UNCLASSIFIED

AD 410261

DEFENSE DOCUMENTATION CENTER

FOR

SCIENTIFIC AND TECHNICAL INFORMATION

CAMERON STATION ALEXANDRIA, VIRGINIA

UNCLASSIFIED

NOTICE: When government or other drawings, specifications or other data are used for any purpose other than in connection with a definitely related government procurement operation, the U. S. Government thereby incurs no responsibility, nor any obligation whatsoever; and the fact that the Government may have formulated, furnished, or in any way supplied the said drawings, specifications, or other data is not to be regarded by implication or otherwise as in any manner licensing the holder or any other person or corporation, or conveying any rights or permission to manufacture, use or sell any patented invention that may in any way be related thereto.

CATALOGED BY DDC
AS AD No. 410261

SEMI-AUTOMATIC IMAGERY SCREENING RESEARCH STUDY AND EXPERIMENTAL INVESTIGATION

Volume I

Report No. 2 and 3

Philco No. V043-2 and 3

Signal Corps Contract No. DA-36-039-SC-90742

U. S. ARMY ELECTRONICS RESEARCH AND DEVELOPMENT LABORATORY
FORT MONMOUTH, NEW JERSEY

Consolidation of Quarterly Report Numbers 2 and 3
1 September 1962 to 28 February 1963

DA Task No. 3A99-23-001-02

410261

DDC
RECEIVED
JUL 17 1963
TISIA D

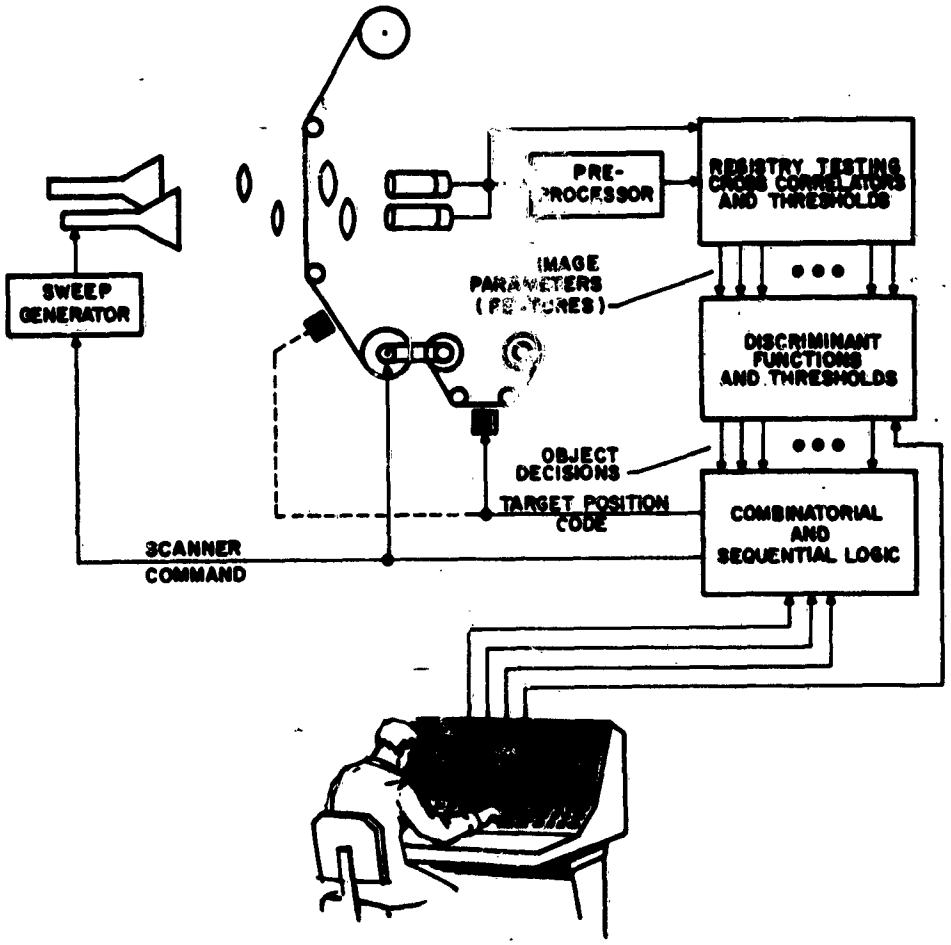
PHILCO
A SUBSIDIARY OF Ford Motor Company
ADVANCED TECHNOLOGY LABORATORY
BLUE BELL, PENNSYLVANIA

NO OTS

ASTIA AVAILABILITY NOTICE

Qualified requestors may obtain copies
of this report from ASTIA

Foreign Announcement and Dissemination
of this report by ASTIA is limited.



Conceptual Block Diagram of Image Screening System

SEMI-AUTOMATIC IMAGERY SCREENING RESEARCH STUDY AND EXPERIMENTAL INVESTIGATION
Volume I

Report No. 2 and 3

Philco No. V043-2 and 3

Signal Corps Contract No. DA-36-039-SC-90742
DA Task No. 3A99-23-001-02
Signal Corps Technical Requirements, SCL-5863 dated 2 October 1961

Dated 29 March 1963

Consolidation of Quarterly Report Numbers 2 and 3

1 September 1962 to 28 February 1963

"The objective of this research study and experimental investigation is to determine techniques and criteria for the later development of a semi-automatic imagery screening device. Screening is defined as 'Gross selection, early in the total interpretation process, to identify those areas in the total supply of imagery which meet the minimum qualification for further interpretation by a human'."

Prepared by:
T. Harley
J. Bryan
L. Kanal
D. Taylor
J. Grayum
H. Kellett

Approved by:

T. J. B. Shanley

TABLE OF CONTENTS

TABLE OF CONTENTS (Cont)

TABLE OF CONTENTS (Cont)

LIST OF ILLUSTRATIONS

LIST OF ILLUSTRATIONS (Cont)

LIST OF ILLUSTRATIONS (Cont)

PURPOSE

"The objective of this research study and experimental investigation is to determine techniques and criteria for the later development of a semi-automatic imagery screening device. Screening is defined as 'Gross selection, early in the total interpretation process, to identify those areas in the total supply of imagery which meet the minimum qualifications for further interpretation by a human'."

The work performed during the reporting interval included the following:

1. Design of a computer simulation experiment to evaluate candidate semi-automatic imagery screening techniques.

2. Investigation of the feasibility of implementation of alternative techniques.

3. Study of the application of statistical decision theory to classification of patterns in reconnaissance imagery.

Philco is doing related work under Contract AF 30(602)-2793, "The Prenormalization of Reconnaissance Data." The objective of that program is the development of techniques for the normalization of target images on aerial reconnaissance photographs before they are used as stimuli in an adaptive-memory type of recognition device.

ABSTRACT

This report covers the work performed under Contract DA-36-039-SC-90742, "Semi-Automatic Imagery Screening," during the six-month period 1 September 1962 through 28 February 1963, comprising the fourth through ninth months of a 21-month program.

The following subjects are presented herein:

(a) Discussion of the merits of available techniques that may be considered for use in each stage of the process of screening aerial reconnaissance photographs for purposes of detecting objects of military significance.

(b) Description and comparison of alternative hardware realisations of the techniques of (a) above.

(c) Summary of the application of the principles of statistical decision theory to the reconnaissance photoscreening problem.

(d) Description of the conceptual design of an imagery screening device.

(e) Detailed description of a computer simulation problem on the detection of tanks in aerial photographs. The simulation, to be carried out during the next six months, is expected to provide design details to extend the conceptual design of item (d).

CONFERENCES HELD WITHIN REPORTING PERIOD

Date	Location	Organizations Represented	Subject and Resume of Conclusions
Oct. 18, '62	Philco, Blue Bell, Pa.	Philco, USAERDL	Program review: Philco presented basic argument in favor of an experimental approach to target recognition system evaluation. Minor changes to be made in Quarterly Report #1 were agreed to by USAERDL and Philco personnel.
Oct. 17, and Oct. 18, 1962	Philco, Blue Bell, Pa.	Philco	Dr. Marvin Minsky, Consultant on this contract, met with cognisent Philco personnel. He made several recommendations which were in general accord with the existing plans.
Oct. 25, 1962	Litton Systems Division, Waltham, Mass.	Philco, USAERDL, Litton	Meeting served as briefing on Government-sponsored work (since concluded) being led by Dr. George Sebastyen on recognition of strategic targets.
Nov. 13, 1962	Evans Area, Ft. Monmouth, N. J.	Philco, USAERDL	Program review: Plans for an experimental program were reviewed and generally approved.
Jan. 22, 1963	Philco, Blue Bell, Pa.	Philco, USAERDL	Program review: Results of implementation study were presented. Reorganisation of Philco laboratories affecting management of the program was outlined.
Jan. 31, 1963	Corning Glass Works, Bradford, Pa.	Philco, Corning	Discussion to determine feasibility of implementing multiple parallel masking logic with optically tapped, acoustic delay line.
Feb. 25, 1963	Evans Area, Ft. Monmouth,	Philco, USAERDL	Presentation of technical data on application of statistical decision theory to automatic target recognition systems.

GLOSSARY OF SYMBOLS

Sections 3 and 4

Symbol	Equation	Definition
$e(t)$	3-2	the noisy input
$e_i(t)$		the value of e(t) at the i^{th} tap of the matched filter at time t
f_j	Figure 3-1	the output of the j^{th} threshold unit
$f(t)$	3-3	the matched-filter output for e(t) input
$f(x, y)$	3-6	the matched filter output
$G(x, y)$	3-9	a thresholded gradient function
$G(\omega)$	3-5	the matched filter function
N_i	3-4	the mean-square noise voltage at the i^{th} tap
$N(\omega)$	3-5	the transform of n(t)
$\lvert N(\omega)\rvert^2$		the power density spectrum of n(t)
$N^*(\omega)$	3-5	the conjugate of $N(\omega)$
$n(t)$	3-2	an additive noise signal
r_i	Figure 3-1	the image signal at the i^{th} input retinal element of a pattern recognition logic
$r_{i,j}$		the value of r(x, y) at the $(i, j)^{th}$ tap of the matched spatial filter
$r(x, y)$	3-6	the noisy input image
$S(i\Delta t)$	3-3	the value of S(t) at the i^{th} sampling point
$S(i\Delta x, j\Delta y)$	3-6	the value of S(x, y) at the $(i, j)^{th}$ sample point

GLOSSARY OF SYMBOLS (Cont)

Symbol	Equation	Definition
$S(t)$	3-1	ideal noise-free input signal to be detected by a matched filter
$s(t)$	3-1	this is a signal defined over all t, which is equal to S(t) during $0 < t < T$, and is zero otherwise
$S(x, y)$		ideal noise-free input image to be detected by a matched spatial filter
$S(\omega)$		the Fourier transform of S(t)
$S^*(\omega)$	3-5	the conjugate of $S(\omega)$
T	3-1	duration of S(t)
t_o	3-2	starting time of S(t)
$v(t)$	4-9	a video signal derived by scanning an image with a spot
w_{ij}	3-7	the weight on r_{ij} in the linear term of a discriminant function
$w_{ij;k\ell}$	3-7	the weight on the product $r_{i,j}\, r_{k,\ell}$ in the quadratic term of a quadratic discriminant function
w_j	Figure 3-1	the weight on the j^{th} threshold output at the final decision summer
$w_{i,j}$	Figure 3-1	the weight on the i^{th} input element at the j^{th} summer
X, Y	3-6	the spatial extent of S(x, y)
Δt	3-3	the Nyquist interval
Δx	3-6	spatial Nyquist interval

Symbol	Equation	Definition
Δy	3-6	spatial Nyquist interval
θ		a threshold value
$\sigma^2\left[e_i(t)\right]$	3-4	variance of the distribution of signal (voltage) values at the i^{th} tap

Section 6

Symbol	Equation	Definition
a_i	6-4 6-5 6-35 6-38	i^{th} coefficient of a linear classification function
a_{ij}	6-7 6-13	ij^{th} coefficient of the quadratic term of a quadratic classification function
b_i	6-7 6-13	i^{th} coefficient of the linear term of a quadratic classification function
C	6-3 6-5	a constant
c_i	6-12	constants arising from the logarithm of the likelihood ratio of joint distributions of binary variables as specified on page 6-7
D^2		Mahalanobis generalized distance
D_N^2	6-24	Mahalanobis generalized distance between two samples based on N measured characteristics
d_j		the difference between the mean values of the j^{th} variable in the two groups

GLOSSARY OF SYMBOLS (Cont)

Symbol	Equation	Definition
E		expectation operator
E_g	6-11	expectation taken with respect to the probability function of group g
F	6-30	the variance ratio of F distribution
F_g		a multivariate probability distribution function for the g^{th} population
f_g		a multivariate probability density function for the g^{th} population
$\hat{f}_g$		the sample estimate of f_g
G		the total number of groups
g		an index taking on values from 1 to G and denoting the group number
i		an index usually running from 1 to N and referring to the i^{th} variable
j		an index usually running from 1 to N and referring to the j^{th} variable
K		a positive integer used in the nonparametric procedure of Fix and Hodges
k	6-10 6-11	an index
k_{jg}	6-14	the cost of classifying group g into group j
L		the likelihood ratio
$M^{(g)}$		mean vector for the g^{th} group

GLOSSARY OF SYMBOLS (Cont)

Symbol	Equation	Definition
$M_i^{(g)}$	6-2	the i^{th} element of $M^{(g)}$; denotes the mean value of the i^{th} variable from the g^{th} population
N		the number of variables measured; the dimension of the space in which samples of patterns are represented as points
n_g		the number of samples from group g
q_g		the a priori probability of group g
$\hat{q}_g$		estimate of q_g
Q_g		estimates for the density functions f_g obtained in the nonparametric procedure of Fix and Hodges
R_1, R_2		decision regions which divide the N dimensional space in which samples are represented
$R_{(j)}$		the expected risk of choosing group j
R	6-31 6-34	Rao's statistic for testing the worth of additional predictors
$r_{ij}^{(g)}$, $r_{ijk}^{(g)}$, $r_{12\cdots n}$	6-11	second, third, and n^{th} order correlation parameters obtained in the expansion of the joint distribution of binary variables for group g
S		sample covariance matrix
S^{-1}	6-15	inverse of sample covariance matrix
s_{ij}	6-18	i, j^{th} element of sample covariance matrix
s^{ij}	6-24	i, j^{th} element of the inverse of the sample covariance matrix

GLOSSARY OF SYMBOLS (Cont)

Symbol	Equation	Definition
t		threshold used in likelihood ratio rule
$u_{ij}^{(g)}$		ratio of the second-order correlation parameter for group g to the product of the standard deviation of the i^{th} and j^{th} variables in group g
V_g		population covariance matrix for group g
V^{-1}		inverse of population covariance matrix
V_{ij}		i, j^{th} element of population covariance matrix
v^{ij}		i, j^{th} element of inverse of population covariance matrix
W		within samples scatter matrix
W^{-1}		inverse of within samples - scatter matrix
w_{ij}	6-27	i, j^{th} element of within-samples scatter matrix
X		vector stochastic variable
x_i		the i^{th} element of the vector X; a predictor variable
X_i^*		the i^{th} element of the selected set of predictors obtained by a screening procedure
x		observed values of the vector variable x
x'		transpose of the column vector x
$\bar{x}^{(g)}$	6-16	the mean of the vector x in group g
$\bar{x}_i^{(g)}$	6-17	the mean of n_g observations on the i^{th} variable x_i in group g

GLOSSARY OF SYMBOLS (Cont)

Symbol	Equation	Definition
$Y_i^{(g)}$		the standardized variable corresponding to x_i obtained by subtracting the mean of x_i in group g from x_i and dividing the result by the standard deviation
Z	6-35	a linear function of the x_i
$\bar{Z}^{(g)}$		the mean value of the samples from group g after projection along a line
π_g		the name for group g
N		Mahalanobis' "true" distance between two populations based on N characteristics

SECTION 1

INTRODUCTION

The principal objective of this program is to complete a broad, comprehensive survey of automatic pattern classification techniques and the alternative implementations of these techniques. The goal of this survey is to obtain insight into the feasibility and usefulness of semi-automatic imagery screening equipment based on these techniques as an aid to the photo interpreter. The work, therefore, is concentrated primarily in two main areas. The first of these is concerned with fundamental processes involved in automatic detection and correlation of militarily significant patterns on aerial photographs. The second area concerns the capabilities and limitations of current and future techniques for carrying out these processes in practical equipment.

The design of a specific semi-automatic screening system is not a goal of the present program. System design is recognized as the next logical step in the overall objective of rendering the current research effort useful to the Army, presuming the prior demonstration of proof-of-feasibility or reasonable probability of success. Consequently, the current program is being pursued in a manner so as to obtain as much preliminary information as possible on the design of a system, consistent with the broader objectives of the general survey.

An analysis described in this report has demonstrated the need for a better and simpler device for the storage and correlation of picture data. A survey of new approaches to meet this need resulted in a concept, also described in this report, for a new type of device that appears capable of overcoming the problem. Another activity in the present program that will facilitate later system design is a brief study, summarized in the present report, of a preliminary conceptual design for a complete target classification system. These examples illustrate the broad survey approach, with design objectives in mind, which characterizes the present program.

SECTION 2

THE IMAGERY SCREENING PROBLEM

In order to carry out a comprehensive investigation of semi-automatic screening techniques, it is essential that the investigator have a general appreciation of the operational imagery screening problem. During this reporting period therefore, Philco project personnel have discussed the problem with Dr. E. A. Rabben of DOD (WSEG), Mr. Lumen of the Army Intelligence Material Development Agency, Dr. Saddacca and Dr. Birnbaum of the Army Personnel Research Office, and have also studied official Army publications[1, 2] related to the subject. A brief summary of the information thus obtained follows.

The photo interpreter normally must detect, identify, locate, and report as many as possible of the militarily significant objects, features and clues in reconnaissance imagery. Sometimes his task is to look for only certain specific information, e.g., nuclear delivery units or armor concentrations.

In the past, each photo interpreter (P.I.) was assigned a particular area of responsibility with which he became thoroughly familiar by studying basic photo cover, maps, and intelligence reports. When new photo coverage was to be obtained, the P.I. team would be informed in advance of the sorties to be flown, coverage to be obtained, scale, and priorities for interpretation. Each P.I. would then collect and study, as time permitted, all available collateral information pertinent to the up-coming screening task. This collateral information was contained mainly in intelligence reports and old photo coverage of the area. When the new photography arrived, it would be divided up, each member of the P.I. team taking the coverage in his area of responsibility. Obviously, this "area of responsibility" technique has its limitations: if most of the coverage fell in the area assigned to one or two men, they would either have to do all the screening that day or accept help from people who were not knowledgeable about the area.

---

1. Department of the Army Field Manual 30-20.

2. Department of the Army Phamphlet 381-1, "Combat Intelligence, Field Army, 1965-1975," September, 1962.

A typical sortie would require each P.I. to go through about 100 photographs. The rapidity with which a man can screen imagery is extremely variable; an expert who is completely familiar with an area may screen most of his pictures in less than a minute each, while less expert P. I.'s who are not familiar with the area may take up to ten minutes per photograph.

The imagery screening task in 1965 and beyond will differ greatly from the typical World War II-Korea task described above. In the first place, a much greater volume of imagery will have to be screened. There probably will be many more photographs per day per unit area than in the past, and most of the photographs will be accompanied by high resolution radar and IR imagery. Future tactical reconnaissance flights will provide three kinds of photo coverage in positive transparency rolls of 5" x 5" frames: panoramic, vertical (1:5000 to 1:12,000), and oblique. In addition, the same flights will produce 5" x 5" frames of high resolution IR imagery at 1:7000 to 1:20,000 scale and 5" x 5" side-looking radar pictures at 1:250,000 to 1:7 million scale.

To be able to screen such a large volume of imagery effectively, the P.I. team will need the best assistance that modern technology can provide in a field environment. The Army is developing the Tactical Imagery Interpretation Facility (TIIF) for this purpose. The first TIIF will contain advanced information-storage and retrieval and display facilities to provide the P.I. with all available collateral information in optimum format. Eventually, the TIIF should evolve into a much more sophisticated system, incorporating computerized collation and analysis of intelligence data, and pattern recognition equipment for screening all types of reconnaissance imagery.

Under the present contract, Philco is directing its attention toward the conceptual design of automatic classification techniques capable of screening high resolution, vertical aerial reconnaissance photographs to locate those areas which most probably contain tactical targets of interest to the photo interpreter. The present phase of the work is directed toward the development of techniques that are capable of recognizing exposed, uncamouflaged targets of a tactical nature, e.g., tanks, vehicles and aircraft (which are clearly recognizable to the human observer), without reference to collateral clues such as tank trackage, location of the possible target on a road or runway, or the nature of the local terrain. This is only one part of the overall imagery screening problem. However, this problem must be solved with relatively simple techniques which can be implemented compactly and economically as a first step toward a useful semi-automatic screening system.

SECTION 3

FUNDAMENTAL CONCEPTS AND TECHNIQUES

3.1 Basic Recognition Concepts

This section presents a description and evaluation of the fundamental processes and techniques involved in the detection of targets of gray-scale imagery. Later sections of the report cover specific technical points relative to the underlying statistical theory as well as the implementation.

In many respects the problem of detecting targets in gray-scale imagery is analogous to the classical communications problem of detection of signal in noise. Both of these problems are amenable to solutions through application of statistical techniques. There are, however, important differences between the two situations. In communication, the problem is usually that of detecting a known signal in a background of additive noise. The noise is usually assumed to be normally distributed in amplitude and independent of frequency (or "white") over that portion of the spectrum of interest. In such a situation, the output signal-to-noise ratio of the receiver is maximized by use of a "matched filter," which may be implemented by a delay line with output taps spaced at the Nyquist interval. The signal outputs from the taps are combined in a weighted summation.

If the ideal input signal to the communications receiver is

$$s(t) = \begin{cases} S(t) & 0 < t < T \\ 0 & \text{otherwise,} \end{cases} \qquad (3\text{-}1)$$

and the actual corrupted signal is

$$e(t) = n(t) + s(t - t_o) \qquad (3\text{-}2)$$

then the output from a matched filter will be

$$f(t) = \sum_{i=o}^{i=\frac{T}{\Delta t}} e(t - i\Delta t)\, S(i\Delta t) \qquad (3\text{-}3)$$

where Δt is the Nyquist interval.

A simple threshold makes the following decisions:

if

$$f(t) > \theta \text{ , then signal present at time t,}$$

and if

$$f(t) < \theta \text{ , then only noise present at time t.}$$

The matched filter concept can be derived from an analysis of linear filter functions which are the same as linear discriminant functions in statistical decision theory. The matched filter is, in fact, a realization of a linear discriminant function operating on the independent random variables $e_i(t) = e(t - i\Delta t)$. Its operation is based on the assumption of normal distribution of the random variables $e_i(t)$ with

$$\sigma^2 [\, e_i(t)] = N_i \,, \tag{3-4}$$

where N_i is the mean squared noise voltage and where the above equation holds for cases of signal plus noise and noise alone.

A more general expression for the matched filter is the function

$$G(\omega) = \frac{S^*(\omega)}{N(\omega)\, N^*(\omega)} \,, \tag{3-5}$$

which is applicable to cases where $|N(\omega)|^2$ is not evenly distributed over the spectrum. This function may have a transform, g(t), which can be implemented in a finite tapped delay line cross-correlator, or it may theoretically require an infinite delay line. In practice, however, implementation of these filters with finite delay line cross-correlators usually yields good approximations to the desired result.

The analogous operation to the matched filter in pattern recognition is the application of a linear discriminant function to the two-dimensional gray-scale signal defining the image. Such a process yields f(x, y) where

$$f(x, y) = \sum_{i=1}^{\frac{x}{\Delta x}} \sum_{j=1}^{\frac{y}{\Delta y}} r(x - i\Delta x,\ y - j\Delta y)\, S\,(i\Delta x,\ j\Delta y) \tag{3-6}$$

where Δx and Δy are the spatial Nyquist sampling intervals, and $S(i\,\Delta x, j\,\Delta y)$ are linear weights. Equation 3-6 describes a pattern recognition technique known as "template matching," which corresponds to making a two-dimensional cross-correlation of the input imagery against an optimized representation of the pattern.

Past experience at Philco in designing recognition equipment, e.g., machine print and hand print readers, has shown that template matching techniques normally are not adequate. The use of the linear discriminant function is optimum only under a very restricted set of conditions. The random variables, $r_{ij} = r(x - i\Delta x,\ y - j\Delta y)$, must be normally distributed and the covariance matrix relating the h_{ij}'s must be the same for both the target and nontarget cases (see Section 6). These conditions do not hold for typical gray-scale inputs or for black and white alphanumerics.

Figure 3-1 is a block diagram representative of a number of suggested pattern recognition systems. These systems all have the common feature that they avoid the limitations of the linear discriminant approach through the use of intermediate sets of non-linear decision elements.

Referring to Figure 3-1, "Simple Layered-Decision Pattern Recognition Logic," many subsets of retinal elements are connected in weighted linear summations as in Equation 3-6, to threshold elements which give unit outputs if the weighted sum exceeds the threshold, and zero otherwise. The outputs from the threshold elements represent a new set of random variables that are connected in a weighted linear summation to a final threshold element that constitutes a final decision.

This basic structure describes feature mapping logics in which the elements in each subset are taken from a small local area of the image, and weighted in a manner that constitutes a matched filter or template for specific sub-portions of the target. The binary random variables from the threshold circuits indicate the presence or absence of specific features in the pattern.

The basic structure described here is used both in the simple perceptrons [1] and by Gamba in his PAPA [2] machine. In these systems, the retinal elements connected to each summer and threshold circuit are selected and weighted randomly within the constraints of the particular model. The selection

1. F. Rosenblatt, Principles of Neurodynamics, Spartan Books, Washington, D. C., 1962.

2. G. Palmieri and R. Sanna, "Automatic Probabilistic Programmer Analyzer for Pattern Recognition," Estro Ho Rivista Methodus, Vol. XII, Number 48, 1960.

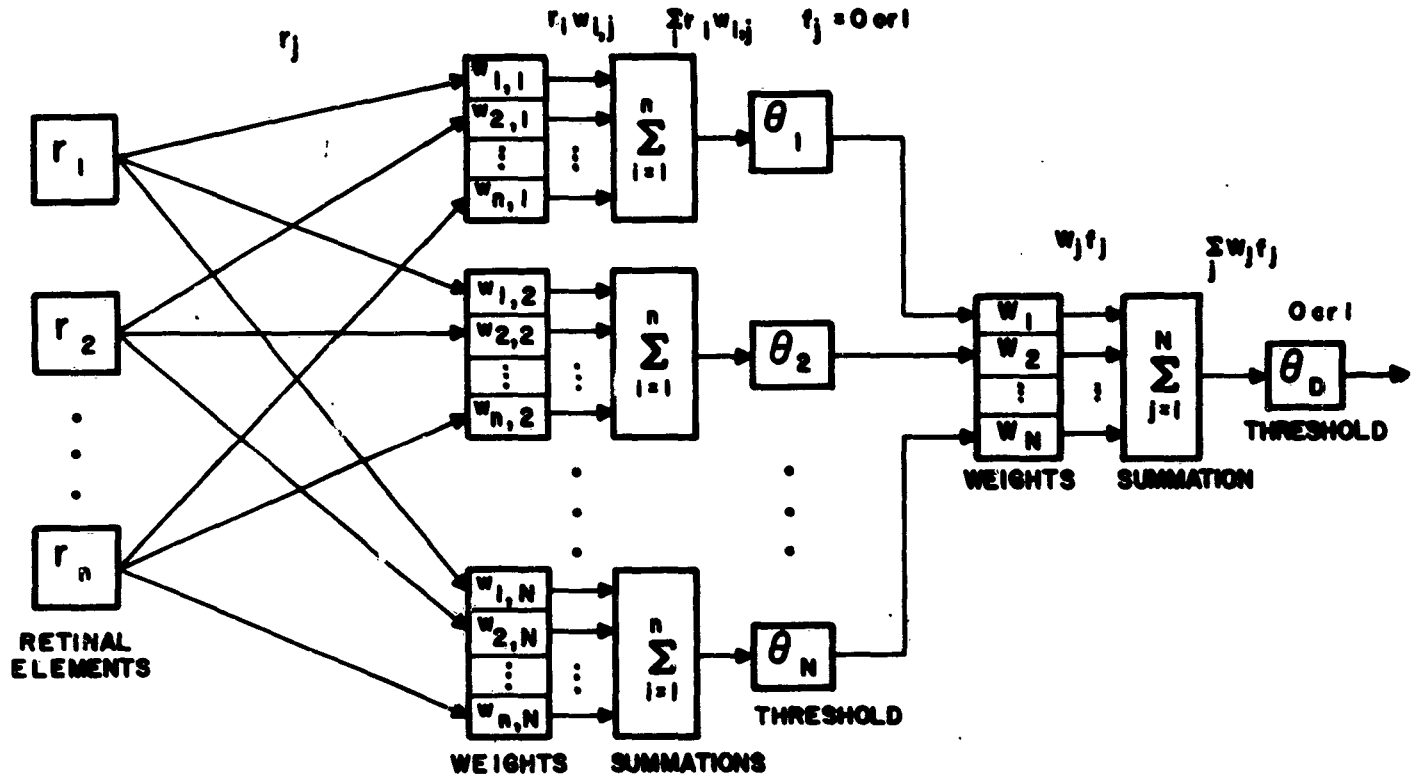


Figure 3-1 Simple Layered Decision Pattern Recognition Logic

of retinal elements purely at random is not an efficient approach to feature selection, and techniques used have varied from random to partially random to completely deterministic wiring schemes depending upon the motivation of the investigator. Gamba, for example, used random-line masks for alpha-numeric recognition in an attempt to take advantage of the connectivity property of printed characters.

In both the "feature map" and the "random mask" systems, the resulting output from the intermediate thresholds is a set of binary random variables. Various adaptive and statistical techniques have been recommended and used to analyze these variables with linear and higher order discriminant functions. Statistical techniques will be reviewed in Section 6 of this report.

The principal property that relates the systems just described is the intermediate set of subdecisions, or thresholds. At first glance the thresholding operation appears undesirable because it destroys information, i.e., it converts an analog voltage amplitude into a binary signal. It is apparent that if the intermediate thresholds, θ_1, θ_2, ..., θ_N, were eliminated from the logic network of Figure 3-1, the resultant network would be equivalent to a single linear discriminant function, or template matching operation, over the entire retinal space. The threshold elements, in association with the linear summations, transform the retinal data, producing a new set of random variables in which the important information is distributed in the marginal probabilities and first-order correlation coefficients where it is more readily accessible to the discriminant function which follows. The specification of threshold elements for this application is not general; other non-linear functions may be applicable in special cases.

The threshold elements, in association with the linear summations, transform the retinal data, thereby producing a new set of binary random variables which are the input variables for another discriminant function. This combination of two layers of linear processing with an intermediate threshold provides a simple and effective way to extract image information that is contained in the interelement relationships and is not accessible to a simple linear discriminant function. The effectiveness of multilayer systems has been demonstrated at Philco in multiple-font print reading equipment.

Another non-linear processing operation that is useful in pattern recognition is detail or boundary detection (see Section 3.2 below). This detection is accomplished by thresholding a spatial derivative (gradient) of image brightness, which transforms the gray-scale picture into a black-and-white "line-drawing" representation which tends to outline objects while reducing the extraneous clutter contained in the original gray-scale image.

Another type of non-linear filtering operation is one which not only adds up weighted sums of retinal element brightness levels as in Equation 3-6, but also takes into account the correlations between brightness levels at different points. This type of operation results in quadratic and higher order discriminant functions that should offer much better performance than simple linear discriminant functions and reduce the number of decision layers needed. A quadratic discriminant function would take the form

$$f(x, y) = \sum_i \sum_j \sum_k \sum_l r_{i,j}\, r_{k,l}\, w_{ij;kl} + \sum_i \sum_j r_{i,j}\, w_{ij} \qquad (3\text{-}7)$$

Implementation of such a function is difficult because the number of product terms in the sum is $n(n-1)/2$ where n is the total number of retinal elements (at least 1000) required to describe a typical pattern. The problem becomes greater with higher order functions. Also, if the input data are not binary, the implementation of the product terms is not simple.

Practical considerations, therefore, limit the amount of higher-order statistical correlation that may be accounted for in recognition logic networks. It is therefore necessary to transform the input random variables -- the retinal element data -- into a new and smaller set of random variables which contain the relevent target information in such a manner so as to be accessible to simple linear, or at most quadratic, discriminant functions. No complete analytic approach to this problem of transforming the retinal data has been discovered. The problem must be solved heuristically with application of statistical techniques wherever appropriate.

In scanning a photograph to recognize a relatively small object of military interest, the photograph can be scanned block by block with overlap between blocks, the individual blocks and the amount of overlap being just large enough to insure that the object will appear within the block in some one position. The object may then appear in any position and in any orientation within the block.

Recognition schemes based on detection of geometrical features must provide for detection of the appropriate features in every possible position and orientation of the object within the block. Testing for different translations at one orientation can be accomplished electronically with one set of feature masks by moving the signals from all resolution elements in the block through a shift register or a tapped delay line. In this report, this procedure is referred to as registry testing. In order to test for unsymmetrical objects at all orientations, it is necessary to provide feature detectors at different angles. For example, if a straight line detector is capable of detecting lines oriented within about ± 12 degrees with respect to a zero line, eight separate line detectors must be provided at intervals of $22\text{-}1/2^{\circ}$. For the detection of translationally or rotationally invariant features such as isotropic texture differences, registry testing and testing for separate orientations are not required.

Using the foregoing basic concepts, the design approach to an image screening system is presented in Tables 3-1 and -2 and in Figure 3-2, "Conceptual Design of Image Screening System." The question arises as to which image characteristics fall into the various classes of parameters, simple objects, and final decisions. To illustrate the general trend, Table 3-3, "Identification of Heavy Antiaircraft Artillery Implacement," has been drawn up from material available in World War II P.I. manuals.[1]

1. TM 30-246, "Tactical Interpretation of Air Photos," Dept. of the Army, February 1954 (also Navaer 10-35-613).

TABLE 3-1

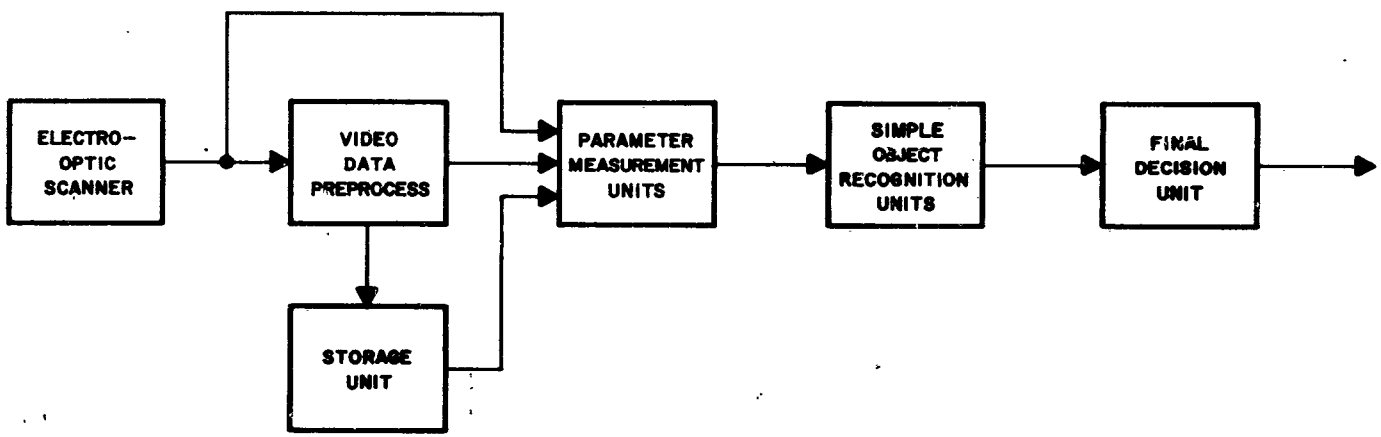
DATA PREPARATION ELEMENTS
OF
IMAGE SCREENING SYSTEM

Unit	Function
Mechanical Transport	Passes images automatically into view of scanner. Probably simple unit for continuous film, more complex for photographs.
Optical or Electro-optical Scanner	Scans images by combination of serial and parallel operations determined by trade-offs between desire for rapid handling versus compact equipment.
Video Data Pre-processor	Normalize video for scale, contrast, etc. Edges detected, silhouetting performed if necessary.
Storage Unit	Storage of parts of image information for correlation of information spanning more than one scan area. Note that use of image as a memory as much as possible will minimize need for high-speed, high-capacity electronic memory.

TABLE 3-2

DECISION ELEMENTS OF IMAGE SCREENING SYSTEM

Unit	Function
Parameter Measurement Units	Extracts detailed information preliminary to and required for the final decision. This preliminary information includes presence of straight lines, intersecting lines, right angles, distinctive texture, etc.
Simple-Object Recognition Unit	From the parameters (see above), decisions are made regarding the presence or absence of simple objects, e.g., tanks, trucks, cable trenches, buildings, etc.
Final Decision Unit	The final decision unit assigns a priority to the photograph or image in terms of the probability of presence of the complex target of interest at the time, e.g., missile sites, camouflaged gun emplacements, etc.



Fi re 3-2 Conceptual Design of Image Screening System

TABLE 3-3

IMAGE SCREENING SYSTEM APPLIED TO LOCATING HEAVY AA EMPLACEMENTS

Identification Characteristic	Functional Unit in Image Screening System
Circular emplacement 16' to 35' diameter	Parameter Measurement Unit (PMU)
Guns in units of 4, 6, or 8 with central CP	Simple Object Recognition Unit (SORU)
Horizontal support girders (for heavy AA guns)	PMU - Straight line detection
Generator Dugout	SORU - Look for characteristic size and irregular shape
Radar and Sound Location Equipment	SORU - May involve complex of parameter measurement outputs
Searchlight Positions	SORU - May involve complex of outputs from PMU
Cable Trenches	PMU - Line detection
Crew Quarters	PMU
Vehicle and personnel paths	SORU

3.2 Preprocessing Techniques

Geometric Normalization

Geometric normalization is a preliminary preprocessing technique for eliminating scale variations in an aerial photograph. It is common practice to record the altitude of the reconnaissance aircraft on each frame. This information can be used directly in conjunction with ground elevation information to compute the necessary adjustments needed to normalize the input to a particular scale factor. This normalization adjustment can be made optically in the sensor portion of the system, or electronically, by adjusting the sweep waveforms of the electronic scanner.

Geometrical normalization is sometimes assumed to include a process whereby objects of interest are centered and placed in a fixed orientation within the frame to be studied. However, before this can be done, the object of interest must at least tentatively be identified and the angular portion of some geometrical features determined. Consequently, it is believed unrealistic to speak of centering and orienting the object prior to identification. These operations, therefore, are not included under the heading of geometrical pre-normalization in the present study. Further, in recognition systems employing element-by-element registry testing, this type of normalization is not necessary because the target is framed automatically.

Detail Detection

Many objects are easily recognized if they are reduced to a line drawing of their principal contours. Techniques that have the property of eliminating the low spatial frequency gray-scale data, and producing a black and white reproduction of the high detail areas including the object contours, are useful. Two candidate techniques have been studied in detail that offer advantages for detail detection: the Laplacian of the image brightness and the Gradient of the image brightness. These techniques, combined with lineness editing to be discussed later, offer a good approximation to the line drawing function. A 'best' technique would serve to reproduce all regularly curving brightness transitions regardless of contrast ratio, and would filter out all other data. As yet, no such technique has been perfected.

The Laplacian is given by the equation

$$\nabla^2 B = \frac{\partial^2 B}{\partial x^2} + \frac{\partial^2 B}{\partial y^2} \qquad (3\text{-}8)$$

where B is the image brightness and x and y are the two image dimensions. Since the function is determined by brightness derivative rather than by brightness itself, it produces no output on large uniform areas. In the presence of line boundaries or corners, the output is high. Thus, the function tends to outline objects and emphasize their contours. It can be simply quantized to two levels without losing outline information. Alternatively, it can be rectified by obtaining $\left|\nabla^2 B\right|$, and then quantizing to produce a passable line drawing of gray-scale image material. Laplacian preprocessing has been used successfully in the Philco Post Office Address Reader. It is identical in function to the two-dimensional detail filter of Taylor,[1] and the convexity detector of Harmon and Van Bergijk.[2] A function similar to the Laplacian has been observed in the frog's retinal ganglion by Lettvin et.al.,[3] and in the lateral geniculate body of the cat by Hubel and Wiesel.[4]

The Laplacian function can be implemented in several ways, both optically and electronically. Optical techniques use coherent optical spatial filtering or a two-channel lensless correlograph. Electronic techniques may utilize special scanning procedures with modulation of the flying spot focus or conventional scanning and lumped constant delay line filters.

The Laplacian function, or spatial second derivative, described above is a scalar quantity derived by a linear process. The Gradient, or spatial first derivative, is a vector quantity which indicates the magnitude and angle of the slope. As such, it cannot be derived by a single linear process; however, simple electronic techniques can be implemented that derive the magnitude of the gradient. This function serves much the same purpose as the Laplacian, but has a better signal-to-noise performance (see Appendix A). When the output is to be thresholded, it is often easier to work with the square of the gradient magnitude, with an output then of the form

$$G(x,y) = \begin{cases} 1 & \text{if } \left|\underline{\nabla} B(x,y)\right|^2 = \left(\frac{\partial B}{\partial x}\right)^2 + \left(\frac{\partial B}{\partial y}\right)^2 \geq \theta \\ -1 & \text{if } \left|\underline{\nabla} B(x,y)\right|^2 = \left(\frac{\partial B}{\partial x}\right)^2 + \left(\frac{\partial B}{\partial y}\right)^2 < \theta \end{cases} \qquad (3\text{-}9)$$

1. Taylor, "Pattern Recognition by Means of Automatic Analog Apparatus," I.E.E., March 1959.
2. L. D. Harman, and W. A. Van Bergeijk, "What Good are Artificial Neurons," Symposium on Bionics, 13-15 September 1960.
3. J. Y. Lettvin et.al., "What the Frog's Eye Tells the Frog's Brain," Proceedings of the IRE, November 1959.
4. Hubel and Wiesel, "Perceptive Fields of Single Neurons in the Cat's Striate Cortex," Journal of Physiology, 1959.

The angle of the gradient vector also may be determined electronically by measuring the horizontal and vertical components of the gradient and computing the arc tan of the ratio. The angular measurements in a local area then can be examined to determine if there is a linear boundary. Such a boundary would result in an array of parallel gradient vectors. This would be true regardless of contrast ratio.

Lineness Editing

Gradient magnitude and Laplacian images produce acceptable line drawings on some input material. Certain types of clutter, however, such as forests, produce strong outputs in the absence of straight lines. If desired, this kind of clutter can be eliminated by subjecting a two-level clipped Laplacian or clipped gradient-magnitude image to a subsequent anisotropic filtering operation. A long, thin aperture is used which is weighted positively in the center and negatively around its periphery, very much like that for the Laplacian except for an elongation in one dimension. The output of this aperture is then presented to a threshold; many such threshold outputs for apertures at eight or more angles of rotation are combined in an OR gate to produce a line drawing output from which clutter has been edited. Figure 3-3 summarizes the method and shows an electronic implementation of the technique.

Silhouetting

A prenormalization technique reported by Holmes, Leland, and Richmond of the Cornell Aeronautical Laboratory,[1] attempts to isolate the object from its background and thus form a silhouette which can be oriented in standard position and subsequently categorized. The technique works well with objects placed on an uncluttered contrasting background, and as one might expect, almost not at all in a highly cluttered background that has approximately the same gray-scale value as the object. The method consists of computing the average brightness over a square "picture frame" aperture, and comparing this to the brightness of the point at the center of the square. A ONE output is recorded whenever the difference of these brightness values exceeds a threshold whose values is proportional to the standard deviation of the brightness values within the aperture. This provides a form of automatic gain control on the decision depending upon the range of brightness values in the area. The output from this prenormalization operation is a binary image in which discrete objects tend to be silhouetted.

---

1. W. S. Holmes, H. R. Leland, G. E. Richmond, "Design of a Photo Interpretation Automation," Proceedings of the 1962 Fall Computer Conference, pp. 27-35.

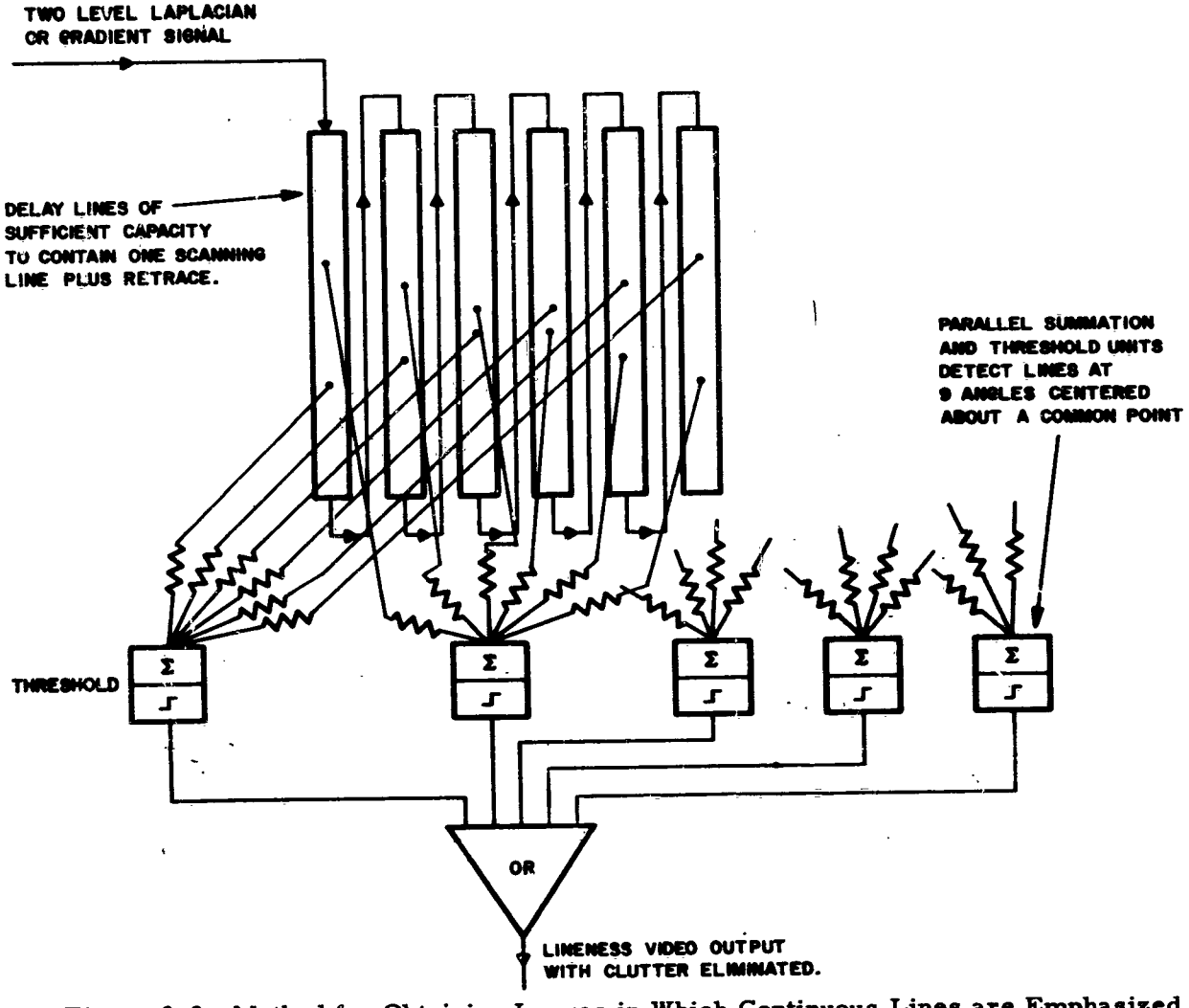


Figure 3-3 Method for Obtaining Images in Which Continuous Lines are Emphasized

Extraction of Additional Information from Gray-Scale Imagery

While the detail detection, lineness editing, and silhouetting techniques are useful in recording fundamental shape information about the image, they may ignore much pertinent information present in the gray-scale data. These processes make use of the high spatial frequency content of the retinal data. The lower frequency retinal data may also supply much information about the nature of the decision area; for instance, oil tanks usually appear to be very bright and of high contrast in respect to their surround. Target recognition devices should utilize all such pertinent information. When images are preprocessed into gradient or lineness images for example, it is important that these other recognition criteria present on the gray-scale image be extracted for later input to the recognition logic rather than be discarded.

3.3 Intermediate Processes - Feature Extraction

General

The function of the parameter measurement unit shown in Figure 3-2 is to perform a number of intermediate transformations on the preprocessed retinal data. The outputs from these transformation operations are a new set of random variables. The transformations are selected to have two primary properties: first, they must continue to preserve the target information, but, second, must present that information in terms of marginal probabilities and joint probabilities of pair states so that the resulting discriminant function has a simple form and may be readily computed. The problem in the selection of the transformation functions is that no comprehensive analytical technique to do this, nor any adequate statistical description of imagery from which analysis might be derived, has yet been developed. Consequently, the selection of good transforms may begin with a list of candidate transforms selected on the basis of some best estimate of utility. The transformations may relate to basic characteristic "features" of the imagery, or they may be selected according to some arbitrary rule or random process. After compiling the list, the problem is to select the good transforms. For this purpose, some analytical tools are available. These will be described in Section 6. The remainder of this section deals with various general classes of transform logics, and how they might be applied to the problem.

There are two approaches to the design of features for incorporation in a target recognition system. The first is to design a universal vocabulary of features which will be generally applicable regardless of what target classes are involved. The advantages are economy of implementation and flexibility of application. If new target classes are added, the only additional circuitry

required is a final decision layer operating on the feature recognition logic outputs. As yet, no adequate set of universal features has been developed. Based on subjective analysis of typical imagery, certain basic features have been suggested, such as straight edges, circles, and other simple geometrical and textural properties. Detailed specification of optimum logics to extract such features has not yet been developed. The use of "random" masks represents another attempt at the generation of a universal vocabulary of "features," based on zoomorphic analogies (modeling of animal processes).

The second approach to feature design is to provide a specific set of features for each pattern class to be recognized. Statistical techniques can be applied effectively in this application. There are many ways in which these techniques may be used. One example is as follows:

The target decision area is divided into a number of feature areas. In each area, the coefficients of a linear or quadratic discriminant function of the form given in Equation 3-7 may be computed from sample values of the input elements for a number of images representing the various pattern classifications. The technique is detailed in a later section. The various methods of calculating coefficients, amenable to programming on a digital computer, are outlined later and detailed in Section 6.

A difficult problem in visual pattern recognition is attaining feature detection which provides adequate tolerance to local translations and/or size and shape variations of patterns within a single class. Possible solutions to this problem include the following.

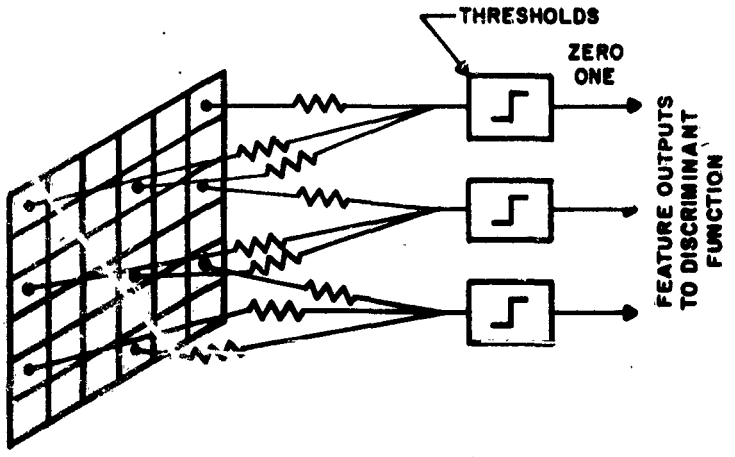
1. Local area threshold logic networks with broad tolerance to the effects of feature translation. However, sufficient tolerance is not always attainable.

2. Use of the approach of Liu and Kamentsky [1] in the process of local area registry testing.

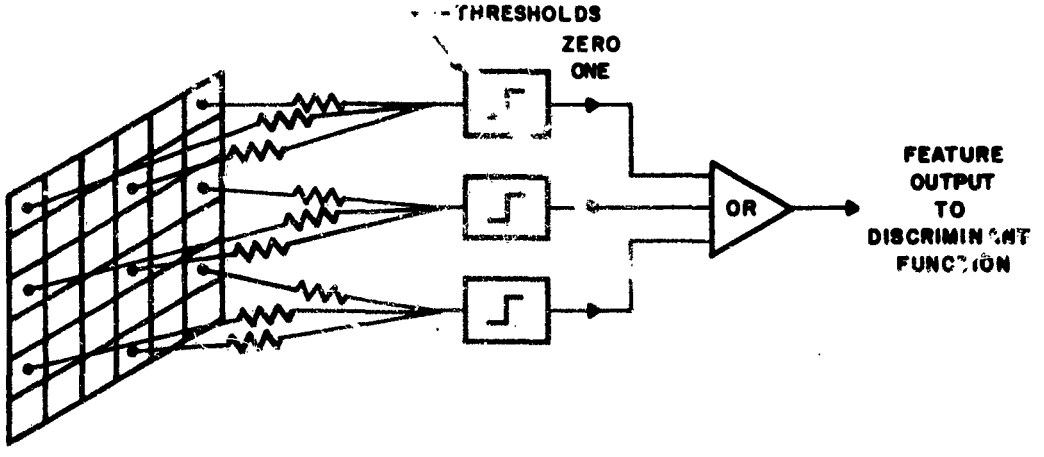
3. Duplicate the feature detection logic in over-lapping fashion to insure feature detection in the presence of translation. Alternative networks to implement this scheme are illustrated in Figure 3-4, "Techniques for the Duplication of Feature Logic."

---

1. L. A. Kamentsky, and C. N. Liu, "Computer-Automated Design of Multi-font Print Recognition Logic," IBM Journal of Research and Development Vol. 7, No. 1, Jan. 1963, pp. 2-13.



(a) IDENTICAL FEATURE DETECTORS DISPLACED IN REGISTRY OVER SMALL LOCAL AREA OF IMAGE WHERE FEATURE IS ANTICIPATED.



(b) FEATURE DETECTORS DISPLACED IN REGISTRY AS ABOVE BUT COMBINED IN OR GATE BEFORE USE IN THE DISCRIMINANT FUNCTION.

Figure 3-4 Techniques for the Duplication of Feature Logic

The basic template matching technique uses a weighted linear summation of the preprocessed retinal data as described by

$$f(x, y) = \sum_{i=1}^{m} \sum_{j=1}^{n} r(x - i\Delta x,\ y - j\Delta y) \cdot w_{ij} \qquad (3\text{-}10)$$

where W_{ij} is the weighting function and $r(x, y)$ represents the preprocessed retinal brightness data. This transformation is the result of the preprocessing operations described in Section 3.2. The output is given by

$$f'(x, y) = \begin{cases} 1 & \text{if } f(x, y) \geq \theta \\ -1 & \text{if } f(x, y) < \theta \end{cases},$$

where θ is a previously established threshold.

For recognition of simple geometric "features" such as straight edges, right angles, or circles, the weights may be chosen arbitrarily to match a particular geometric configuration. For example, the recognition of horizontal straight edges could be based on a linear template applied to thresholded gradient magnitude data, where the template has the following form.

j \ i	1	2	3	4	5	6	7	8	9	10
1	-w	-w	-w	-w	-w	-w	-w	-w	-w	-w
2	+w	+w	+w	+w	+w	+w	+w	+w	+w	+w
3	+w	+w	+w	+w	+w	+w	+w	+w	+w	+w
4	-w	-w	-w	-w	-w	-w	-w	-w	-w	-w

Then

$$f_{horiz}(x, y) = w \left[\sum_{i=1}^{10} [G(x - i\Delta x,\ y - 2\Delta y) + G(x - i\Delta x,\ y - 3\Delta y) \right.$$

$$\left. - G(x - i\Delta x,\ y - \Delta y) - G(x - i\Delta x,\ y - 4\Delta y) \right] \qquad (3\text{-}11)$$

where $f_{horiz}(x, y)$ can range in values from +40 w to -40 w. An area of uniform brightness, or a contrast edge not approximately parallel to the straight line mask will result in $f_{horiz}(x, y) = 0$. Therefore, for a threshold well above zero but somewhat less than 40, 25 for example, this template will be a reasonably good straight edge segment detector. The output from such a feature template matching unit f(x, y) will be an array of plus ones (corresponding to a detection) and minus ones, (no detection) covering every position (x, y) in which the template was registered.

The output from any location (x_1, y_1) will be a particular sample of a random variable $f(x_1, y_1)$ having a high probability that $f(x_1, y_1) = 1$ if an edge of a man-made object is present there, and -1 if it is natural terrain. The spatial distribution of f(x, y) over the decision area will be effective in determining whether a specific class is present.

This template matching logic is identical in structure to the A-units of the Perceptron. The usual procedure is to generate an output from each A unit for only one position of registry within the decision area, but this is not a fundamental limitation. That is, each random mask can be matched to the pre-processed retinal data in many (or every) position of registry. The principal characteristic of the A-units in Perceptron application is that the weights and connections to the retinal element are not designed to correlate to "features" but are chosen according to some random process. This process can set various constraints, such as "choose points that are contiguous along a curving line," and still represent a random selection procedure. The random variables, $x_i = +1$ or -1, outputs from the A-units, having no overt functional or semantic relationship to the input data, are presumed to be uncorrelated, or at least correlated similarly (equal covariance matrices) in the signal plus noise (target) and pure noise (no target) cases. Under these assumptions, the optimum decision logic is a linear discriminant function of the form $\Sigma a_i x_i$ whose coefficients may be calculated from the responses of the masks to a sample set of images. The determination of the weighting functions, a_i, whether carried out by statistical or adaptive procedures, provides a selection process for the random masks. Those that respond fairly uniformly to both the targets and

non-targets in the training sample will have very low weights and will be effectively eliminated. The masks that respond most often to only one of the two classes in the sample set will have high weights and will be retained. The principal limitation of these masks is that the amount of information contributed by any relatively small set of them may be inadequate for good target recognition, and further, that as additional A-units are used, the cross-correlation coefficients may be such that they contribute no new information to a decision based on a linear discriminant function.

The spatial filtering operation described above can be applied either in the spatial frequency domain by multiplying the signal transforms using a spatial frequency filter, or in the space domain through a cross-correlation of a weighting function with the transformed or raw retinal data in every position of alignment (see Appendix B). Selection of a preferred technique can be made solely on the basis of feasibility and cost of implementation. There are several equivalent hardware implementations for such a filtering operation. Optical techniques, e.g., lensless correlographs and coherent light optical spatial filters, provide parallel access to an area of the input image and can perform the cross-correlation of a single weighting function simultaneously over all possible alignments. Electronic techniques based on scanning of the input image provide sequential access to the points of the image, and must be used in conjunction with a dynamic memory such as a delay line or shift register. With such a memory many weighting functions can simultaneously be cross-correlated at any given moment in a single relative alignment with the image, with all possible alignments or position of registry of the image being accomplished time sequentially.

Digital Design Techniques

The preceding discussion has emphasized the use of threshold logic to detect individual features to be combined for a subsequent decision. Threshold logic is attractive in pattern recognition because it can interpolate between samples where only one or two bits or picture elements may be different between samples. Conventional parallel digital logic, however, is often easier to implement and service. With digital logic, once a satisfactory design has been achieved, there are no problems of threshold drift and adjustment.

On the other hand, digital logic does not have any inherent ability to interpolate from known samples to unknown or unspecified samples. Digital logic works only to the extent that the designer or design algorithm is able to foresee and provide for all possible logical alternatives.

Little progress has been made over the last few years in designing general-purpose algorithms to produce digital logic which extrapolates from a

training sequence to provide for unknowns. Carne, working with the Artron at Melpar,[1] and Powers and Vernot at Philco,[2] have attempted to design general-purpose adaptive logic for pattern recognition application. In each case, the result has been a device which can memorize a training sequence perfectly, but is unable to do much better than chance on unknowns. Algorithms have been designed successfully at Philco to handle digital design for specific problems, such as recognition of hand-print alphanumeric characters, but each new problem has required a new algorithm.

One approach to digital design specifies the problem by the use of a truth table (see Figure 3-5). Once the truth table is specified, it is a relatively simple matter to implement it with digital logic. The chief problem with this technique is that if N digital inputs are available, a truth table has 2^N entries which must be filled to specify the problem. Usually an insufficient training sample is available to fill all entries.

Decision tree techniques for pattern recognition are well known. Basically, a decision tree consists of a series of dichotomous decisions placed on a sorting path as shown in Figure 3-5b. At each decision, the path forks into two alternatives. A sample for test enters the tree at the top, and after passing a series of tests ends up at a terminal classification. Each junction usually represents the testing of a particular bit in an input code. A well designed sequential decision tree tests only those bits necessary for classification; the remaining bits are ignored. A decision tree, however, need not represent a sequential process. All tests may be performed simultaneously and their outputs combined by parallel logic to achieve the same terminal classification. Of course, all tests may not be used in every classification. Feigenbaum and Simon [3,4] have programed a general-purpose algorithm for the design of a decision tree from a training sequence. The routine is part of a program called EPAM (Elementary Perceiver and Memorizer). The EPAM program has been used thus far to associate written and verbal representations

---

1. E. B. Carne, E. M. Connolly, P. H. Halpern, and B. A. Logan, "A Self-Organizing Binary Logical Element," Biological Prototypes and Synthetic Systems, pp. 311-330, Plenum Press, N. Y.

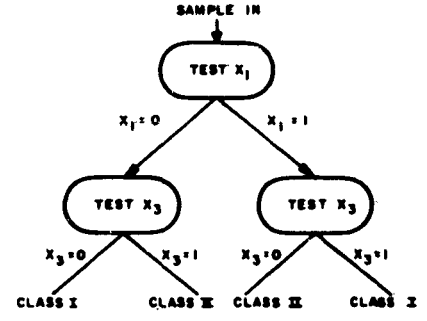
2. R. D. Vernot and E. N. Powers, "A Tunnel-Diode Adaptive Logic Net," Proceedings of the 1962 International Solid State Conference.

3. E. A. Figenbaum, and H. A. Simon, "Performance of a Reading Task by an Elementary Perceiving and Memorizing Program," RAND Report, p. 2358.

4. E. A. Figenbaum, "The Simulation of Verbal Learning Behavior," Proceedings of the Western Joint Computer Conference, Vol. 19, 1961, pp. 121-132.

X_1	X_2	X_3	CLASS
0	0	0	I
0	0	1	II
0	1	0	I
0	1	1	II
1	0	0	II
1	0	1	I
1	1	0	II
1	1	1	I

(a) Example of a classification problem expressed by a three variable truth table in which classification for all states is specified by an exhaustive training sequence for three input variables, X_1, X_2, and X_3

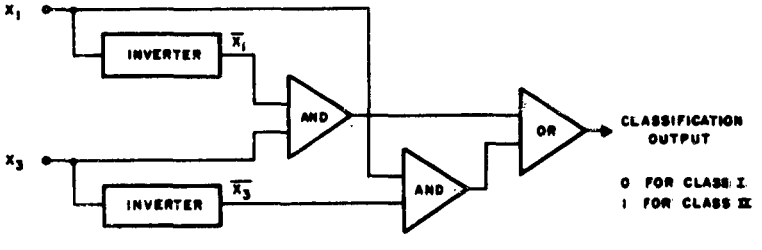


(b) Decision tree to classify truth table entries designed by Feigenbaum's algorithm. Tree discloses that input variable X_2 is not needed for classification. Logical relation expressed by decision tree

$$I = X_1 \cdot X_3 + X_1 \cdot X_3$$

$$II = X_1 \cdot \overline{X}_3 + \overline{X}_1 \cdot X_3$$

where . indicates AND
+ indicates OR
and — indicates NOT



(c) Resulting logical design using parallel logic

Figure 3-5 Digital Design Approach to Feature Extraction

of nonsense syllables. EPAM has been observed to make mistakes similar to those made by humans on similar problems. No data are available however on its ability to generalize to unknown samples.

Some investigators have avoided the problem of adaptive digital design by means of constrained random generation of digital combinations of binary input variables. The first constraint chosen is the number of input variables to be considered at any one time in logical combinations. Thus, for a particular number choice, n, a series of n-tuples are generated. When n equals 2, the logical states of binary picture-element pairs are considered, when n equals 3, triplets, and so forth.

Once the n-tuples are generated, their digital outputs are tabulated over a training sequence and weighted by some rule (usually statistical in nature) to be combined in a subsequent decision. These techniques are close relatives of perceptron-like devices and to the kind of threshold feature detectors considered in this report, except that the randomly chosen n-picture elements are combined digitally rather than by threshold logic. If a separate weight set and threshold were provided for each n-tuple state recorded, if all weights were set to plus or minus ones, and if the thresholds were set to the highest possible value, then the two techniques become identical.

Browning and Bledsoe [1,2] experimented with n-tuples, with n having values as high as six. In these works, all states of the n-tuple were recorded and used; thus, for n = 6, 2^6 or 64 states were recorded for each n-tuple. Simple Bayes' weighting was used and the results were excellent for machine print and encouraging for handprint. Uhr and Vossler [3] have designed a program which records only one state each of many 25-tuples. The program incorporates rules for discarding the least useful digital combinations and generating new ones. This program has demonstrated some success in recognising cartoon characters and recognizing speech in binary sonogram representation.

---

1. W. W. Bledsoe, and C. L. Bison, Improved Memory Matrices for the n-Tuple Pattern Recognition Method, internal report, Panoramic Research, Inc.

2. W. W. Bledsoe, and I. Browning, "Pattern Recognition and Reading by Machine," Proceedings of the Eastern Joint Computer Conf., 1959.

3. L. Uhr, and C. Vossler, "A Pattern Recognition Program That Generates, Evaluates, and Adjusts its Own Operators," Proc. Western Joint Computer Conf., 1961, pp. 555-569.

Promising work in n-tuples has recently been published by Kamentsky and Liu, [1] who have expanded on the n-tuple concept in multiple-font print-reading application. The n-tuples of Kamentsky and Liu differ from those of Browning and Bledsoe in that registry testing and memory have been added to the n-bit states. The Kamentsky n-tuple is a constrained random choice of seven binary picture elements in the input image. These seven elements are given ZERO and ONE designations and a ONE output is recorded for the n-tuple if all seven ZERO-ONE input states are satisfied simultaneously in any position of translational registry over the input character. This registry testing aspect is quite important since it increases the probability of a ONE response by an order of magnitude, thus increasing the power of the n-tuple to transmit information. The choice of the 7-tuple over other numbers was based on experiments which determined the information content of n-tuples for various values of n. The results of Kamentsky and Liu are sufficiently impressive for the alphanumeric problem to merit further investigation of n-tuples in object recognition applications. Points which need further investigation are:

1. Relative efficiency of logical n-tuples versus threshold n-tuples. The logical AND used by Kamentsky to combine the elements of his n-tuples is equivalent to a simple linear summation of the elements with a threshold on the output set to fire only when all of the elements are active. Therefore, this point of investigation, in more general terms, is to determine how the efficiency of a threshold n-tuple varies as a function of its threshold.

2. Optimum constraints for n-tuple generation.

3. Efficiency of n-tuples versus the amount of registry testing used. Local area registry testing of a digital n-tuple increases the probability of a ONE response and so increases its ability to transmit information. On the other hand, further registry testing over a wider area causes information loss, to the extent that relative feature position may be important to recognition. The relative trade-offs of these considerations need to be investigated.

Most logical processes can in theory, be represented by parallel logic networks. However, as a practical matter, many useful recognition routines are impractical to implement with parallel logic because of the excessive number

1. L. A. Kamentsky, and C. N. Liu, "Computer-Automated Design of Multi-font Print Recognition Logic," IBM Journal of Research and Development, Vol. 7, No. 1, Jan. 1963, pp. 2-13.

of components required. Complex image processing techniques often can be implemented most economically by some form of sequential logic. Much of the initial work in pattern recognition used sequential logic.[1,2] Routines were programed for the general-purpose digital computer which smoothed edges, eliminated clutter, and traced contours to extract major geometric features from alphanumeric characters. Work at Cornell Aeronautical Laboratories[3] in the isolation of silhouettes from gray-scale imagery is an example of sequential routines applied to gray-scale object recognition. Sequential routines, of course, are relatively slow. Parallel and sequential techniques may be combined as a compromise between speed and equipment complexity. For example, it may be desirable to obtain tentative identifications of objects using parallel logic, and subsequently to verify this identification with a more sophisticated sequential logic operation. Contour following is an important example of a sequential processing technique. There is little doubt that human eye motions follow contours in tracing paths through clutter.[4] Contour following may be an efficient technique for final recognition after the number of candidate patterns has been reduced by high-speed parallel preprocessing.

Translationally Invariant Feature Extraction Techniques

Much attention in the pattern recognition art has been given to the role of two features of the image which are invariant to translation of the input. These are the Fourier energy spectrum and the autocorrelation function. These two are very closely related, the first being the Fourier transform of the latter. They are similar also in that they destroy the same information in eliminating the effects of translation. The following discussion is concerned with the Fourier energy spectrum only. It should be understood, however, that completely analogous statements can be made about the autocorrelation function because of the transform relationship existing between the two functions.

---

1. J. S. Bomba, "Alpha-Numeric Character Recognition Using Local Operations," Proceedings of the Eastern Joint Computer Conference, 1959.

2. S. H. Unger, "Pattern Detection and Recognition," Proc. of the IRE, October 1959.

3. W. S. Holmes, H. R. Leland, G. E. Richmond, "Design of a Photo Interpretation Automation," Proceedings of the 1962 Fall Computer Conference, pp. 27-35.

4. J. R. Platt, "How We See Straight Lines," Scientific American, Vol. 22, No. 6, pp. 121-129, 1960.

The Two-Dimensional Fourier Transforms

The expression

$$R(\omega_x, \omega_y) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} r(x, y)\, e^{-j(\omega_x x + \omega_y y)}\, dx\, dy \qquad (3\text{-}12)$$

retains all of the information of the input image. From the transform output $R(\omega_x, \omega_y)$, one can, using the inverse transform, reconstruct the original image $r(x, y)$ with no loss of information, including the exact position of the image with respect to the origin. The image data are contained in both the amplitude and the phase of the transform. To obtain an output which is not a function of input translation, something must be discarded. The information of image location with respect to the origin is contained exclusively in the phase data. Thus, when the Fourier energy spectrum $R(\omega_x, \omega_y) \cdot R^*(\omega_x, \omega_y)$ is obtained, the output is independent of input image translation. Unfortunately, more is lost than just total image translation. This single parameter might, for example, be specified by 12 bits: six bits in x and six in y. However, in discarding the phase data, about half of the information in the transform is lost; in a picture involving 64 x 64 elements resolution, 3-bits each, this might involve as much as 12,000 bits. The power spectrum no longer specifies the image completely, and it is impossible to reconstruct the image from the power spectrum alone. In fact, more than one image might well have the same power spectrum. Figure 3-6 shows an example of two 1-dimensional functions which have the same power spectrum.

Figure 3-7 shows a one-dimensional Fourier spectrum plot for a tank in a cluttered background. Here the spectrum plot at each vertical position corresponds to the spectrum of a corresponding narrow horizontal strip taken across the picture at the same height. Sharp peaks can be observed by the fundamental frequency and harmonics of the periodic bogie suspension. Also, a single peak can be observed at the height of the upper tread, corresponding to the frequency of the tread links. Figure 3-8 shows two-dimensional Fourier energy spectra of a white rectangle on a black field, and a tank photograph. The spectrum of the rectangular aperture is very nearly ideal with the complete bilateral symmetry which is to be expected with a pure, real input function $R(\omega) = R(-\omega)$. In the case of the tank, however, the spectrum lacks the expected symmetry, indicating that variable emulsion thickness of the input transparency has caused variable phase shift even with the negative immersed between optical flats in an oil having approximately the same index of refraction. The photographs shown in Figures 3-7 and 3-8 were produced in the Philco laboratories

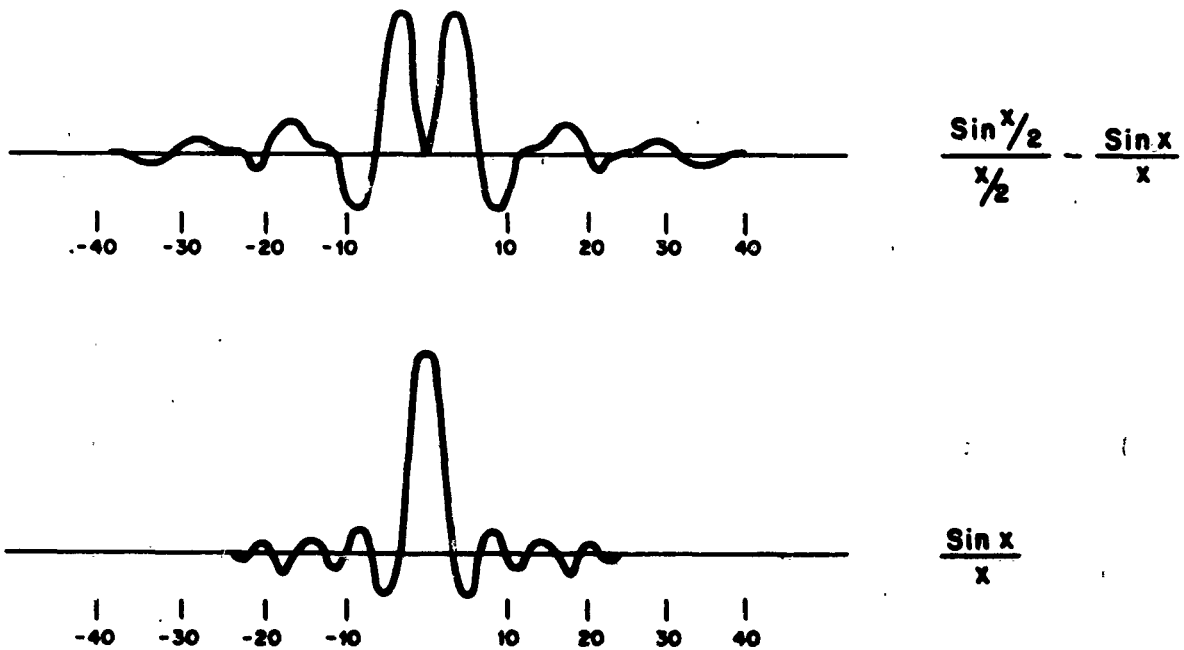


Figure 3-6 Two One-Dimensional Functions Having Identical Fourier Power Spectra

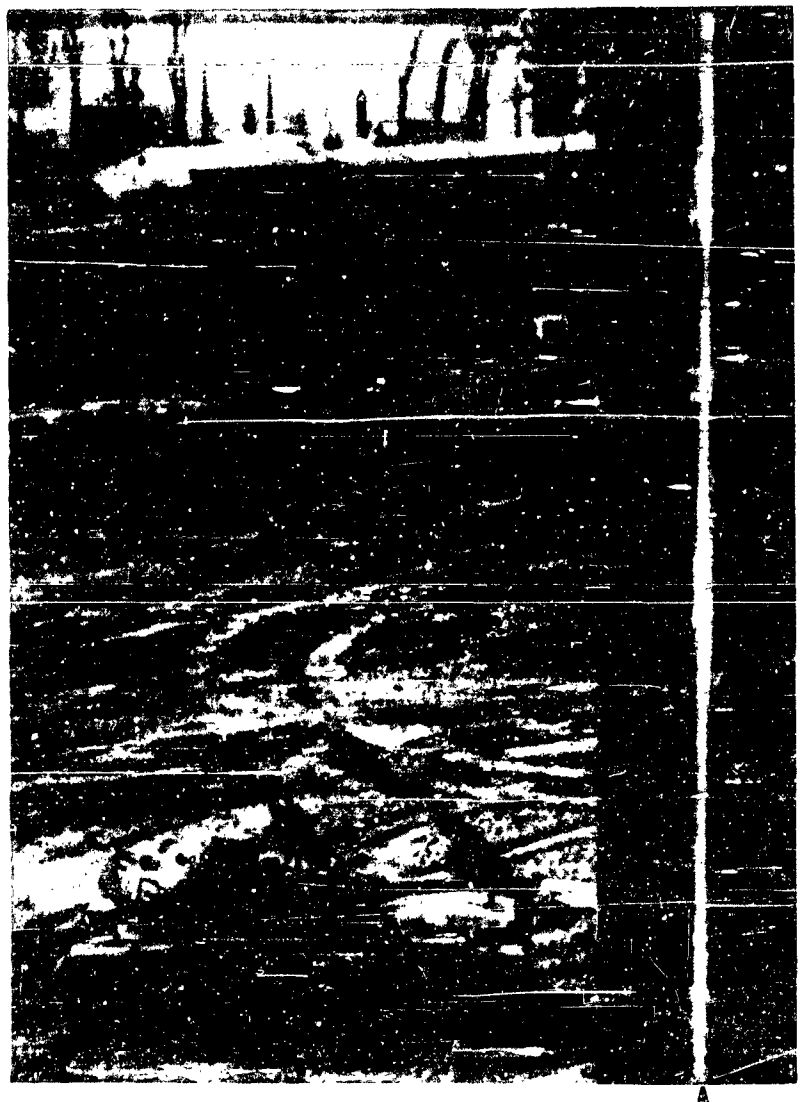
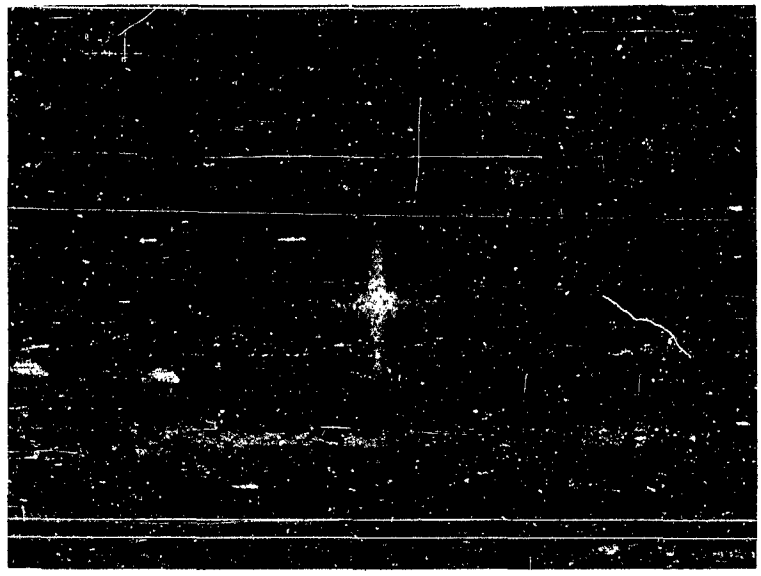


Figure 3-7 One-Dimensional Spectrum of Tank and Background Scene



(a) Spectrum of a Rectangular Aperture

(b) Spectrum of Tank Photo (Showing Film Thickness Effects)

Figure 3-8 Two-Dimensional Spectra Obtained with Mirror Spectrometer

in connection with another contract. Figures 3-9 and 3-10 show the optical set-ups used. Many more photographs would be required to determine whether or not Fourier spectra contain sufficient data for dependable recognition of various classes of objects.

The problem of large area texture identification appears to be well adapted to the use of Fourier power series techniques. Since no geometric relationships need be preserved in the application, the discarded phase information is of lesser importance.

If either the Fourier power spectrum or the autocorrelation function is to be useful to eliminate translation, the total frame size over which the spectrum is evaluated must be larger than the object to be recognized. If, for example, they were the same size, the object would be already centered and there would be no need for a translationally invariant processing technique. Unfortunately, as the frame size is expanded to include more than the target, clutter in the target background adds to and dilutes the target spectrum making recognition more difficult.

Textural features are recognized by applying linear discriminant functions to the components of the Fourier power spectrum. The output ϕ_i for the textural feature T_i, is the weighted summation of all the components of the power spectrum. The component weights can be derived from sample data in the same manner that the coefficients of any linear discriminant function. In Appendix C, it is demonstrated that this is completely equivalent to cross-correlating the input material, h(x, y) with each of two weighting functions $w_1(x, y)$ and $w_2(x, y)$, squaring the outputs, and integrating the difference over the frame.

$$f_1(x, y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} r(x - \xi, y - \eta)\, w_1(\xi, \eta)\, d\xi\, d\eta \tag{3-13}$$

$$f_2(x, y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} r(x - \xi, y - \eta)\, w_2(\xi, \eta)\, d\xi\, d\eta \tag{3-14}$$

$$\phi_i = 4\pi^2 \iint_{\text{Frame}} \left| f_1^2(x, y) - f_2^2(x, y) \right| dx\, dy . \tag{3-15}$$

Figure 3-9 One-Dimensional Multichannel Spatial Spectrometer

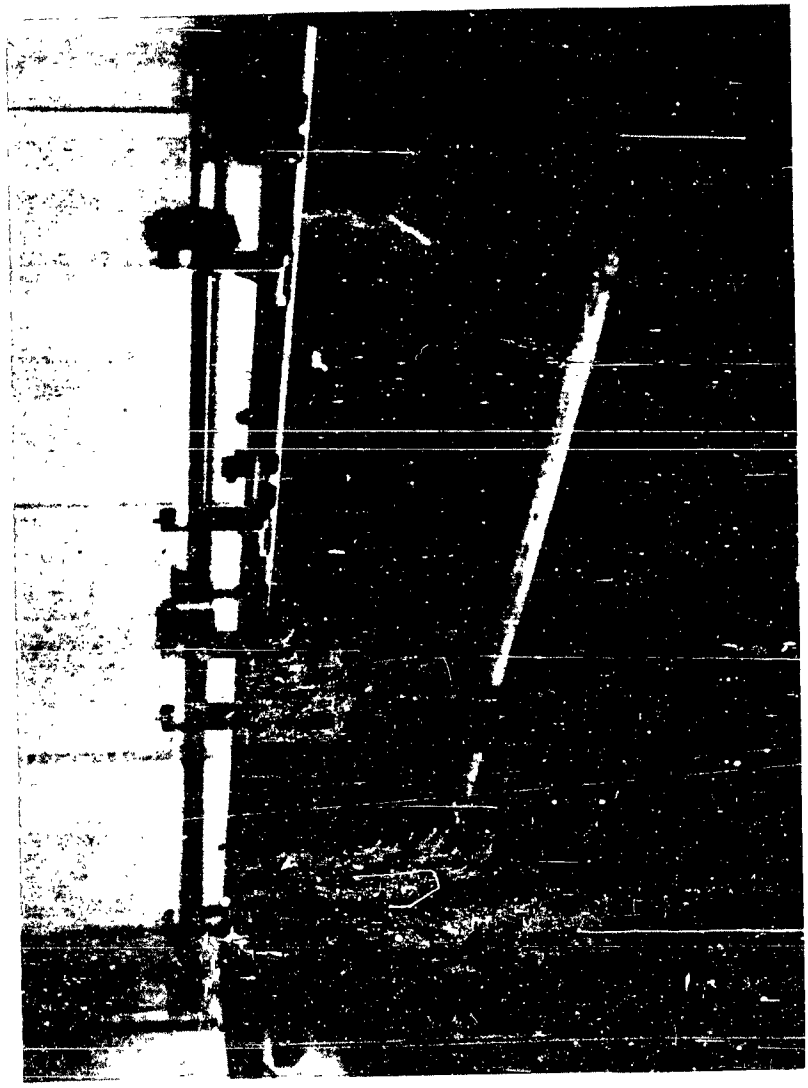


Figure 3-10 Spatial Spectrometer

To make a texture measurement in a small local area, it is only necessary to reduce the area of frame. The resulting output is no longer a rigorous equivalent of the weighted Fourier power spectrum measurement, however, the loss of rigor is no worse than that associated with any other method which measures Fourier spectrum over a small area.

The textural measurements, then, can be made either electronically or optically by the same basic techniques used for geometric feature extraction but with the addition of a squaring operation and an integration or summation over the frame.

3.4 The Role of Statistical Decision Theory

The intelligent application of techniques derived from statistical decision theory is basic to the solution of the imagery screening problem. Specific applications have been touched upon briefly in earlier portions of Section 3.

Statistical decision theory is applicable to the following type of problems.

- Unknown samples are drawn from a universe made up of several classes.
- A set of measurements are made of properties of each unknown sample. Based on these measurements, the sample is to be classified.
- For each class in the universe, the multivariate probability distribution of these property measurements is known, or can be estimated from a set of known samples.
- The "costs" of misclassifying a sample of a certain class into each of the other classes is known or estimated.
- The distribution of each class within the universe is known or estimated.

Under these circumstances, statistical decision theory can provide useful rules for classifying the unknown samples. The specific rules to be used will depend on the nature of the multivariate distributions of the property measurements and the completeness of the description of the universe. Techniques available fall into two general classes: parametric, in which the parameters * of the distributions are known or assumed; and nonparametric, in which

* e.g., the moments of the distributions

no assumption is made about the distributions. The principal techniques are described in Section 6 of this report. The usefulness of any given technique will depend on the specific application.

In the processing of gray-scale imagery for target detection, statistical techniques may be applied in a number of ways. Each element of the retinal field of gray-scale or two-level edge detected values constitutes a random variable. If complete statistics were known describing the multivariate distribution of the retinal element values for each target class and for all non-target images, then, at least theoretically if not practically, target detection could be performed optimally by a single layer of logic operating on the retinal data. The linear "matched filter" or template matching logic described earlier is an extremely simple single-layer logic, based on very restrictive assumptions about the distribution of the retinal element data.

In multilayer systems, the output of the feature detector constitutes a set of random variables. Because of the practical problems of implementation, useful features will be those that present the information necessary for classification in terms of much simpler statistical distributions. In Section 7 of this report, an experimental program is described for investigating the application of statistical techniques operating first on the retinal layer to extract features, and then on the feature detector outputs to make the final classification.

SECTION 4

SYSTEM IMPLEMENTATION TECHNIQUES

4.1 General

There are two general approaches to the implementation of an image screening system. The first, based principally on optical techniques, provides parallel access to a large area of the input imagery for cross-correlation in all positions of registry with one weighting function or template at a time. The second approach, making use of image scanning and delay line or shift register cross-correlators, uses sequential access to positions on the retina, but provides for simultaneous cross-correlation of many templates.

The investigation and comparison of the sequential versus the parallel approach is still in progress. The comparison must necessarily be a detailed one since the investigation must consider the design details and the hardware implementation involved. The present material, therefore, reports the current status of a continuing program and is in no sense a final evaluation.

4.2 Parallel-Access Sequential-Masking

Introduction

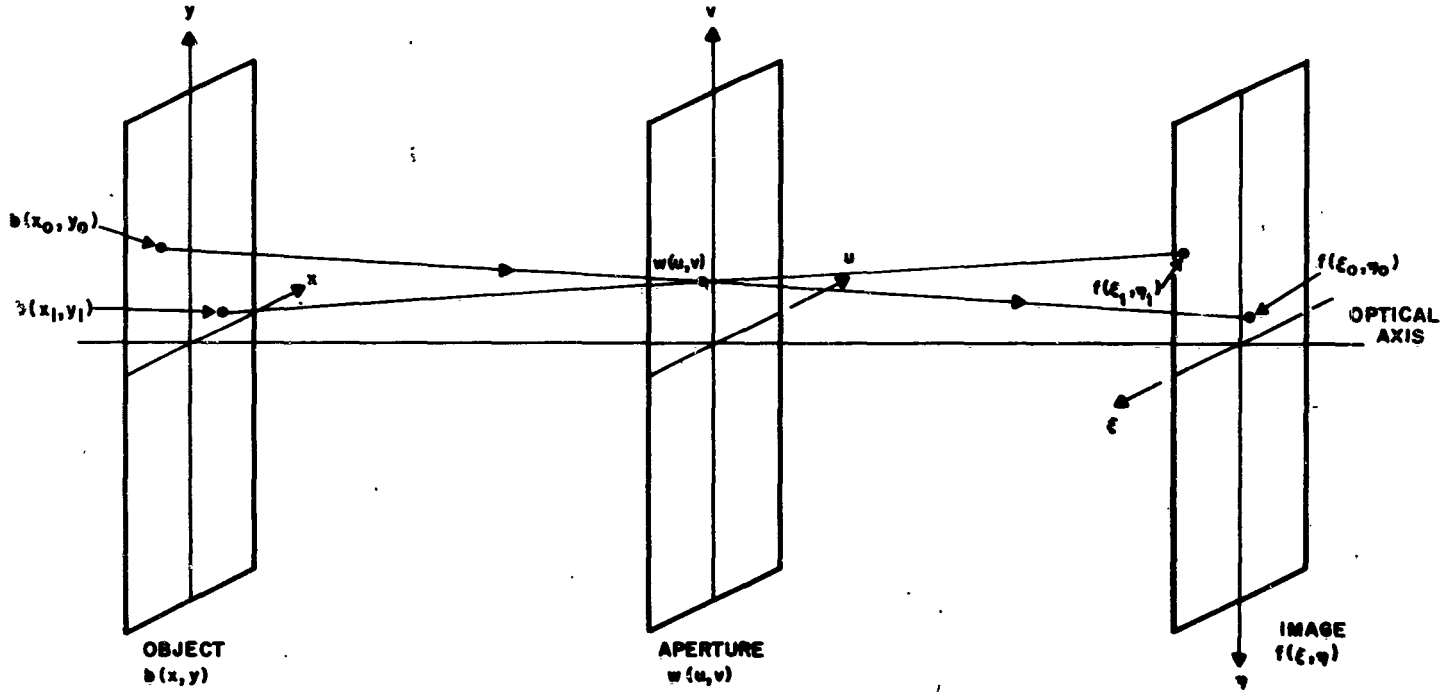
Two principal techniques -- the lensless correlograph and coherent optical spatial filtering -- provide for simultaneous correlation of a single template or filter function in every position of registration with the input image. Both have been described in detail in the literature* as single-layer, linear template-matching devices. Their effectiveness in multilayered systems is presently limited by the available sensors.

The Lensless Correlograph

The basic operation of the lensless correlograph (Figure 4-1 a and b) can be understood in terms of a generalized pinhole camera, with the pinhole replaced by a shaped, weighted aperture. The weights are achieved by controlling the transparency at each point of the aperture in accordance with weighting function. The aperture may be a continuous function but the operation of the correlograph may be best understood in terms of

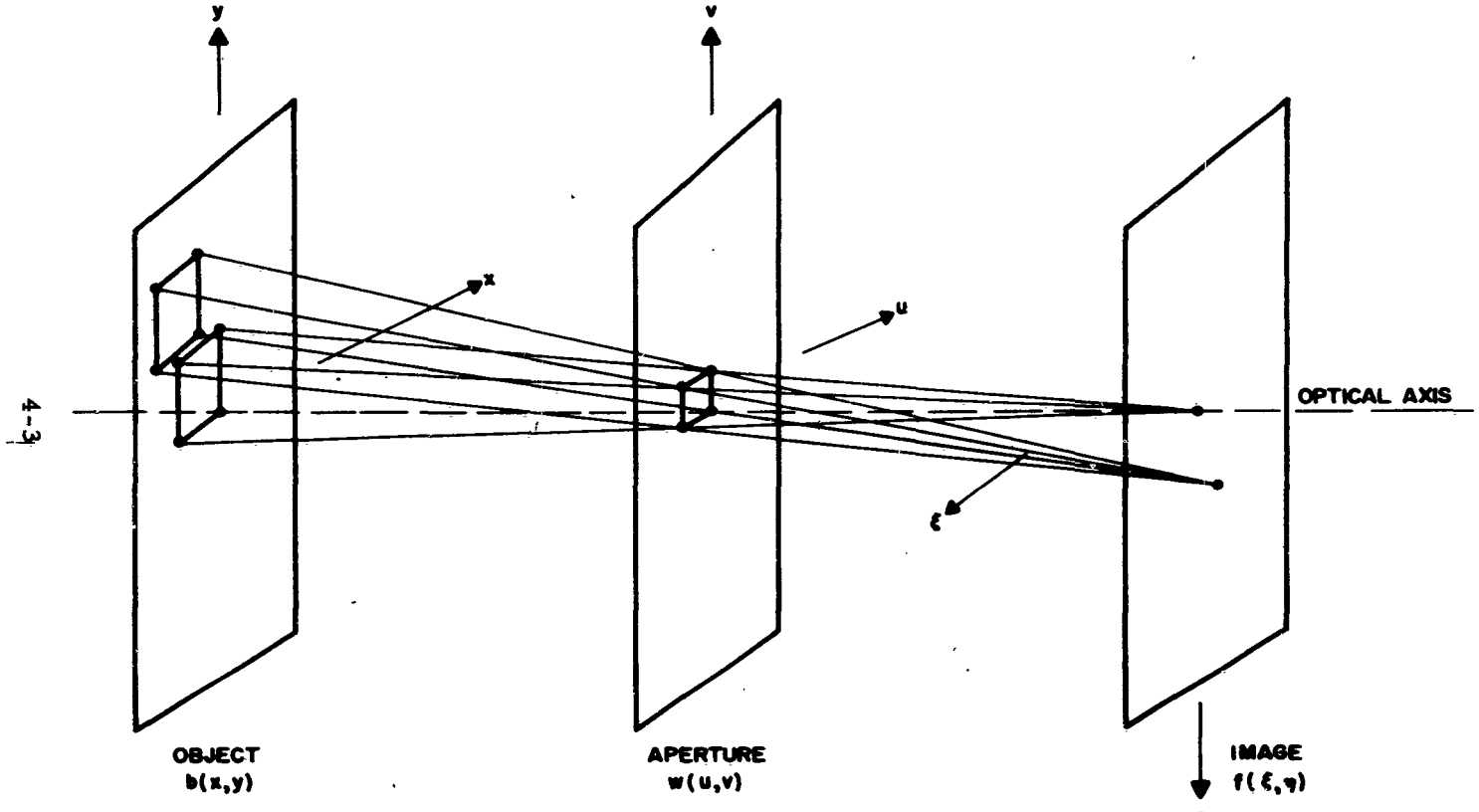
---

* See bibliography in First Quarterly Report.



$$f(\xi,\eta) = b(\xi-2u, \eta-2v)\, w(u,v) = b(x,y)\, w\left(\frac{\xi-x}{2}, \frac{\eta-y}{2}\right)$$

Figure 4-1(a) The Lensless Correlograph



$$f(\xi,\eta) = \iint b(\xi-2u,\eta-2v)\,w(u,v)\,du\,dv$$

Figure 4-1(b) Lensless Correlograph - Shaped Aperture

discrete cells, or picture elements. For a single cell, the operation is identical to that of a pinhole camera. Each point of the input photograph, b(x, y) in the object plane is imaged at a point, (ξ, η) on the image plane that is complementary to it with respect to the aperture cell at (u, v). For a correlograph with unity magnification, as in Figure 4-1a, the elementary output $f_e(\xi, \eta)$ on the image plane is defined as

$$f_e(\xi, \eta) = b(x, y)\, w\left(\frac{\xi - x}{2}, \frac{\eta - y}{2}\right) \tag{4-1}$$

or

$$f_e(\xi, \eta) = b(\xi - 2u, \eta - 2v)\, w(u, v) \tag{4-2}$$

where

w(u, v) is the weighting value at point (u, v) .

For the case of a shaped weighted aperture, as in Figure 4-1(b), the output is the integral over the aperture of $f_e(\xi, \eta)$:

$$f(\xi, \eta) = \iint_{\text{Aperture}} b(\xi - 2u, \eta - 2v)\, w(u, v)\, du\, dv \tag{4-3}$$

or

$$f(\xi, \eta) = \iint_{\substack{\text{projection of aperture on x, y object plane from point } (\xi, \eta)}} b(x, y)\, w\left(\frac{\xi - x}{2}, \frac{\eta - y}{2}\right) dx\, dy \tag{4-4}$$

The resulting output is a cross-correlation of the aperture with the input photograph in every position of registry.

In this implementation, $w(u, v)$ is a positive real function, and the output $f(\xi, \eta)$ i. also positive real. It is impossible to implement directly negative lobes on the weighting function. "No correlation" corresponds to a positive dc value rather than to zero.

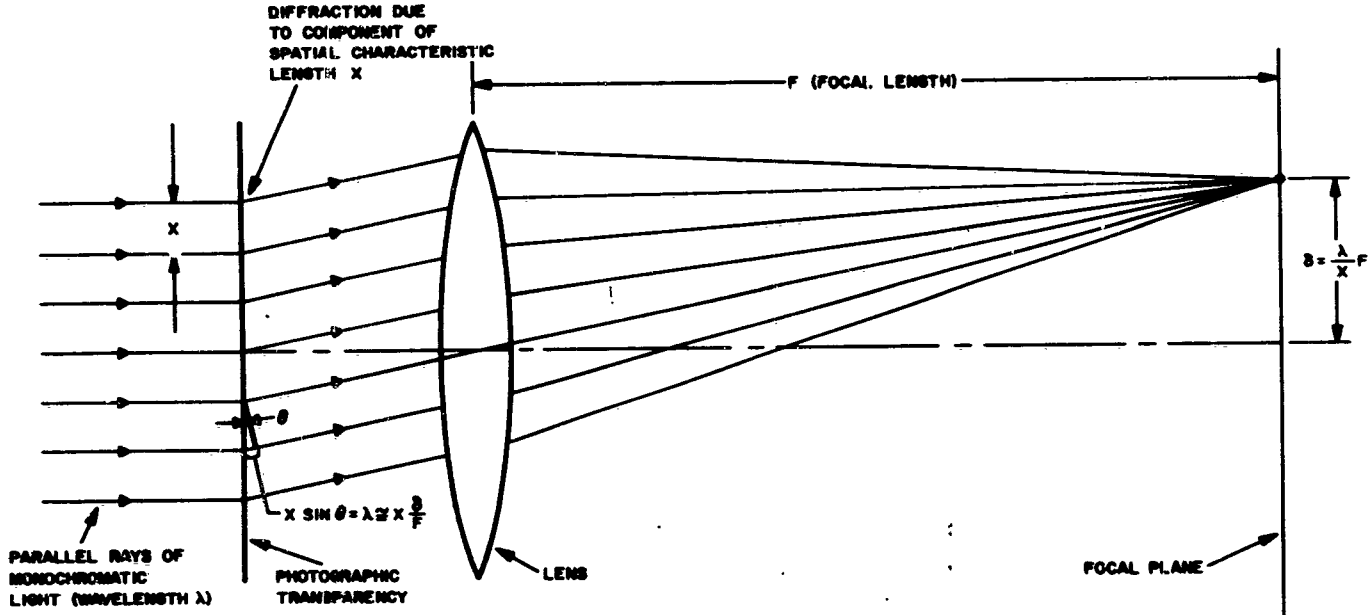
The lensless correlograph produces a bright correlation peak in light intensity at diagonally opposite points in the image plane. This peak corresponds to the location of a feature in the object plane for which a correlation filter is placed in the aperture plane. In order to detect specific single objects or features, a thresholding operation must be performed at each point in the output image. Object counting corresponds to integrating the thresholded output over a frame.

The practical advantages of the lensless correlograph include its relative simplicity and its ability to test for features independent of position of those features in the image field, all in one parallel operation. Disadvantages include the inability to perform non-linear transformations (e. g., gradient magnitude) in a direct way, difficulty in obtaining adequate light levels in the output, difficulty in obtaining fast, efficient, stable, and uniform parallel thresholding logic sensors to use in the output stage, and inability to incorporate negative weights in the aperture.

Coherent Optical Spatial Filtering

The coherent optical spatial filtering technique provides spatial filtering directly in the spatial frequency domain. When a collimated beam of monochromatic light is passed through a transparency and a high quality lens, under ideal conditions, the distribution of light in the focal plane of the lens corresponds to the Fourier spatial frequency transform of the transparency. As the parallel rays of light pass through the transparency, they are diffracted or bent at each point in proportion to the spatial frequency content at that point. In the focal plane, all the unbent (dc component) rays converge to a point on the optical axis. All the rays bent by a given angle, corresponding to a given spatial frequency, also emerge in parallel from the transparency and are converged by the lens to a single off-axis point in the focal plane, with the distance from the optical axis to the point proportional to the magnitude of the corresponding spatial frequency (see Figure 4-2). This property can be used for measurement of the textural parameters of the imagery by sensing various components of the power spectrum.

When a second lens is added to the system, the input transparency is imaged on the "film" plane. The Fourier spectrum in the focal plane is



NOTE:
GENERATION OF FOURIER SPECTRUM IN FOCAL PLANE (X DIRECTION ONLY)
COORDINATE OF POINT CORRESPONDING TO COMPONENT IN TRANSPARENCY OF
SPATIAL WAVELENGTH X IS GIVEN BY $\delta = \frac{\lambda}{X} F$.

Figure 4-2 Generation of Fourier Spectrum in Focal Plane

inversely transformed into a reproduction of the input image. However, when a spatial filter in the form of a transparency is placed in the focal plane, the image reproduced is the filtered version of the input image (see Figure 4-3). That is,

$$f(x, y) = \mathcal{F}^{-1}\left[B(\omega_x, \omega_y) \cdot G(\omega_x, \omega_y)\right] = b(x, y) * g(x, y), \qquad (4\text{-}5)$$

where $f(x, y)$ is the output image, $b(x, y)$ the input image and $B(\omega_x, \omega_y)$ its transform, $G(\omega_x, \omega_y)$ the spatial filter function and $g(x, y)$ its inverse transform. The resulting function $f(x, y)$ could also have been derived with a lensless correlograph or electronically, using scanning and a delay line cross-correlator.

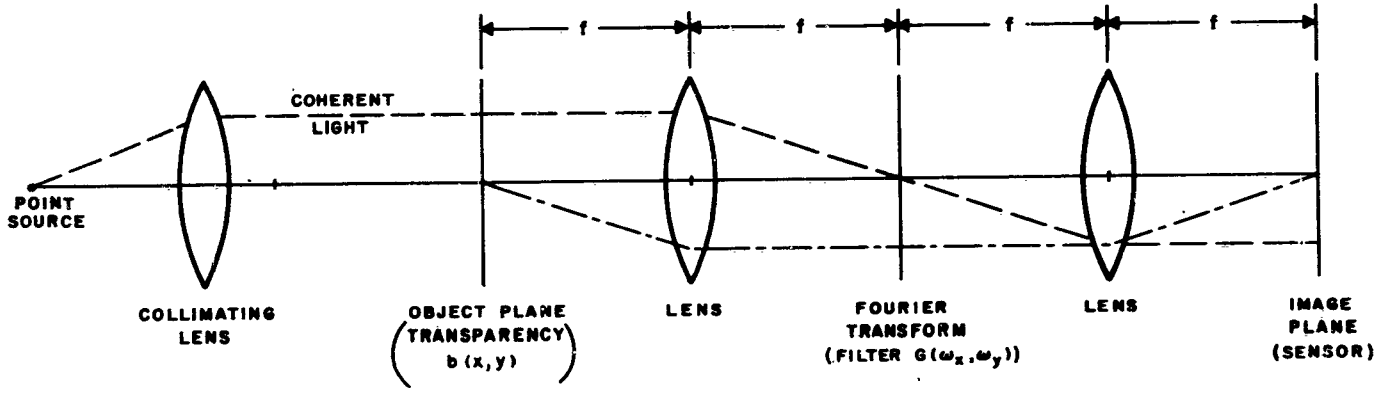
The coherent light optical spatial filter technique appears to have the same basic advantages and disadvantages as the lensless correlograph listed in the previous paragraph, except that it can incorporate negative weights in the aperture. Other practical problems arise in obtaining an adequate coherent light source and a good optical system. The implementation of a desired filter function may be difficult also, since the filter must vary in both transmission (amplitude) and thickness (phase). Phase variations can be obtained by varying the thickness of the filter, but such filters are difficult to make.

The output of a lensless correlograph, or coherent optical spatial filtering system, is a two-dimensional filtered image of the input transparency. The principal advantage of optical filtering is that all the filtered data is available simultaneously. The parallel sensor must be capable of thresholding the signal on a point-by-point basis over the image. Solid state light amplifiers offer promise as sensors for the application. The use of film is another possibility, although the processing involved is a distinct drawback.

Solid State Optical Decision Panels

The basic components of the solid state optical decision panels (see Figure 4-4) are a sandwich of transparent conductor, photoconductive layer, non-linear resistance layer, electroluminescent layer, and a second transparent conductor.

The photoconductive (PC) layer is a semiconductor whose conductance (1/R) is approximately proportional to the intensity of light impingent



NOTE: IMAGE AT SENSOR IS DEFINED BY $b(x,y) * g(x,y)$
WHERE $g(x,y) = \mathcal{F}^{-1}\left[G(\omega_x, \omega_y)\right]$

Figure 4-3 Coherent Light Optical Spatial Filtering System for Unity Magnification

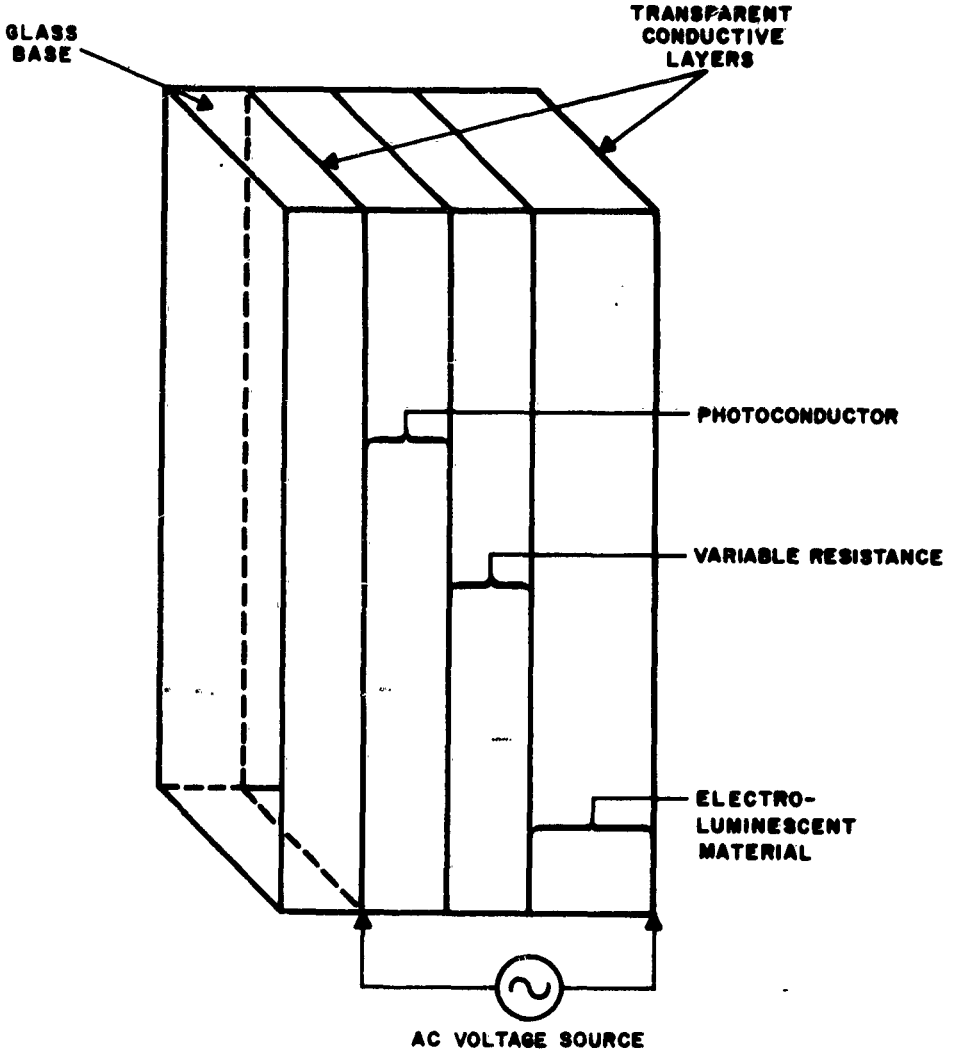


Figure 4-4 Schematic Diagram of Solid State Optical Decision Panel

upon its surface. Resistance changes over a range of 4000 to 1 are not uncommon with cadmium sulfide and cadmium selenide, both of which are commonly used in photoconductive panels. The non-linear resistance is silicon carbide, commonly used for making varistors. The electroluminescent layer, (EL), emits light when an ac potential is applied on the layer. The light output is a function of the amplitude and shape of the applied voltage waveform. The output depends on the material used, but the general form of the output has been observed to be

$$B = A(\omega) \exp (b V^{1/2}) \text{ ft-lamberts} , \qquad (4\text{-}6)$$

where $A(\omega)$ is a characteristic constant for a particular ac frequency ω, b is a constant, the value depending upon the material, and V is the amplitude of the ac voltage across the EL layer.

The sandwich functions as a highly non-linear light amplifier, and can be used as a thresholding array. An ac voltage is applied across the sandwich, between the two transparent conductive layers. When no light hits the panel, the resistance of the photoconductor is high and little current flows. As a result, the varistor-type layer is of high resistance, thus further lowering the current. The voltage across the EL layer is very low, and no light is radiated.

When light strikes a point on the panel, it lowers the photoconductor resistance, resulting in increased current. This decreases the varistor-layer resistance, and the voltage across the EL layer increases sharply. This in turn results in emission of light at all points where the light input exceeds a particular threshold value. If the impinging light was a high-pass filtered image of a target, the light output would be a detail detected reproduction of the target.

The speed of a solid-state decision panel is determined by the light flux density striking the EL layer. The average illuminance (flux density) in the image plane of a lensless correlograph, where the panel might serve as a sensor, is given by Equation 4-7, which is derived in Appendix D.

$$\bar{E} = \frac{m}{n} \frac{B}{64\pi} \left(\frac{a_o}{a_o + a_i} \right)^2 \qquad \frac{\text{lumens}}{\text{square foot}} \qquad (4\text{-}7)$$

where m is the number of elements in the aperture,

n is the total number of picture elements in the input

B is the peak value (white) of luminous emittance in the input in lumens/square foot, and

a_o and a_i are the linear dimensions of a picture element in the object and image plane respectively given in any consistant units. The values of a_o and a_i are determined by the input photograph resolution and output sensor resolution, respectively.

Possible data rates based on the correlograph in combination with the EL-Varistor-PC panel can be computed assuming some typical values for the parameters. Using a 10-element aperture with a tungsten wire source at 3200°K, over a portion of the input image 1000 elements square having 2-mil resolution and a sensor with 10-mil elements, we obtain

$$m = 10$$

$$n = 10^6$$

$$a_o = 2 \text{ mils}$$

$$a_i = 10 \text{ mils}$$

$$B = 10^6 \text{ lumens/ft}^2$$

$$\overline{E} = \frac{10 \cdot 10^6}{10^6 \cdot 64\pi} \left[\frac{2}{10+2}\right]^2 \qquad (4\text{-}8)$$

$$\overline{E} = 1.4\,(10)^{-3} \text{ lumens/ft}^2 .$$

The decision time (rise plus decay) for one aperture under the foregoing conditions is approximately 0.1 second. By using a larger aperture, times on the order of 0.02 second are expected. These figures are based on characteristics of an EL-PC sandwich designed for linear operation, i.e., without a silicon carbide layer.

The basic idea of a solid-state, optical decision panel is very attractive because of the possibilities it offers for high-speed, parallel data processing.

The principal limitations of state-of-the-art solid state light amplifier panels are:

1. Light outputs are too low to be useful as input to succeeding layers of optical logic (≈ 10 ft-lamberts).

2. Stable PC layers with uniform photoconductivity over the entire retinal space are very difficult to obtain.

3. Resolution is limited by the EL layer to about 0.01 inch per element which is over ten times as large as a film resolution element.

4. Decisions cannot be stored directly for later processing in conjunction with other data.

Film Techniques

High-resolution, rapid-process film may be used to record the output data and to preserve the information after development for additional parallel processing. Currently available films have optical resolutions of better than 1000 lines/mm, a figure which is more than required for recognition by current methods. The exposure time for each element under the same conditions as for the EL panel would be approximately 2 seconds. Processing speed is 5-10 seconds; however, as many frames as necessary may be processed simultaneously, so that the exposure time, rather than film processing time, is the limiting factor.

Film has the following advantages:

1. Ease of procurement and reliability

2. Permanent storage of information

3. Non-linearities can be realized in the film, i.e., the film can perform thresholding.

The disadvantages are:

1. Additional development of hardware is necessary to enhance film operation

2. Operating conditions must be carefully controlled (temperature, humidity, etc.).

Sequential Detection After Optical Filtering

The signal at the output of the optical systems can be detected sequentially using one of the scanning devices described in the next section; however, this technique does not take advantage of the parallel access property of optically filtered images. In addition to its utility for feature detection, as described in the foregoing subsections, optical filtering of an entire image may be useful for simple preprocessing operations preliminary to sequential electronic scanning, e.g., where a high-pass spatial frequency filter is applied optically to the input image for edge enhancement.

4.3 Sequential-Access, Multiple Parallel-Masking Techniques

Scanning Techniques

Electronic scanning techniques are of course not limited to the conventional TV scan. For target recognition systems the raster generating waveforms can be adapted to the specific needs of the logic. For applications where it is desired to cross-correlate a particular weighting function in all positions of registry, the optimum raster waveform is a sawtooth ribbon scan, with height greater than, or equal to, the template height (see Figure 4-5). The time-varying waveform output from a ribbon scan n elements wide is related to the retinal elements by the equation

$$v(t - i\Delta t) = r\left[x - (i \bmod n)\,\Delta x,\ y - \frac{(i - i \bmod n)}{n}\,\Delta y\right] \quad i = 0, 1, \ldots, \tag{4-9}$$

where t is a time reference corresponding to point (x, y) on the retina, and Δt is the temporal Nyquist interval corresponding to the spatial interval, Δx (= Δy). A similar equation can be written for a continuously varying signal, v (T + t), in terms of r (X + x, Y + y), but is more complex since it involves continuous variation along x, but discrete steps Δy along y corresponding to the interval between scanning lines.

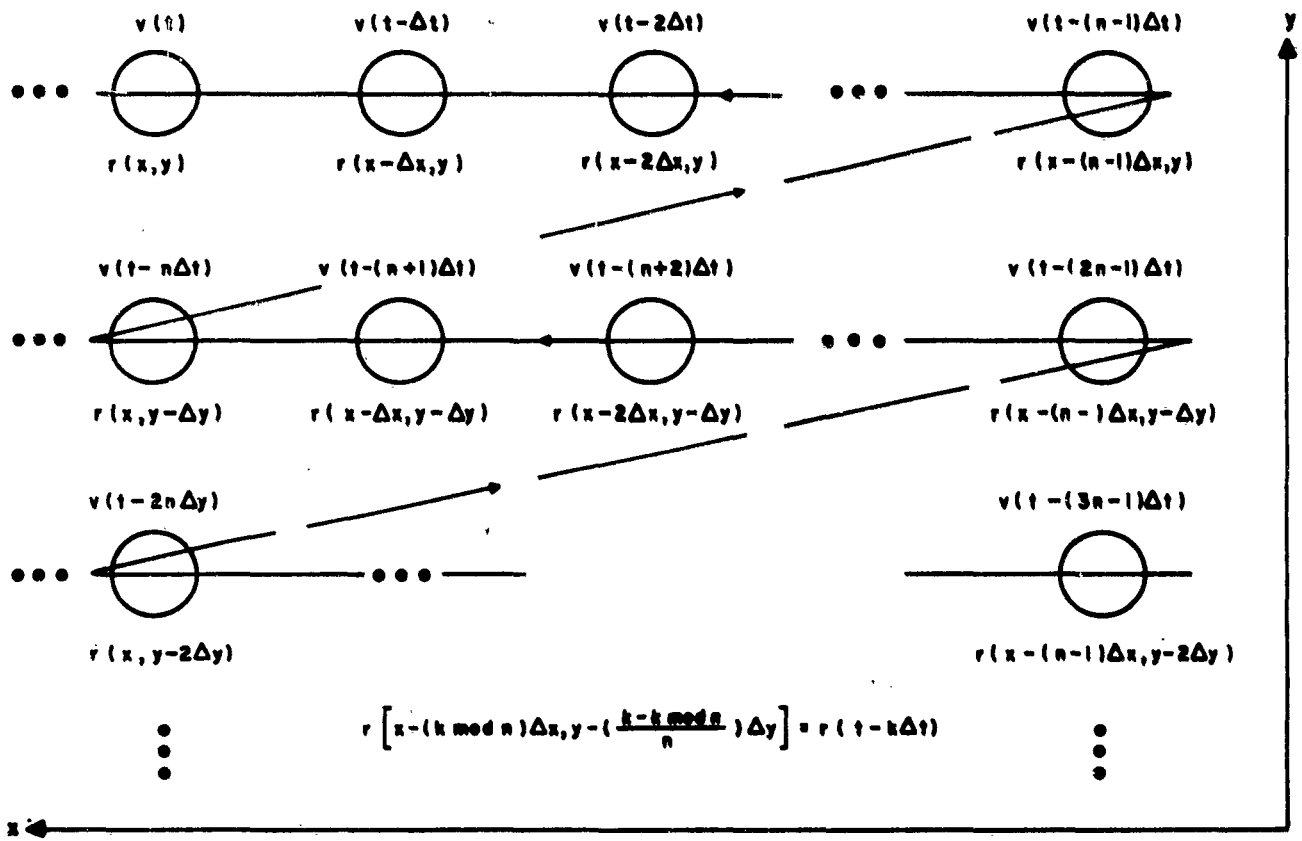


Figure 4-5 Relationship Between Retinal Elements and Ribbon Scan Video Output Elements

The delay line cross-correlator serves as dynamic storage for the input video, $v(t)$. The output at any tap, $v_k(t)$, is identical to the input video with a fixed time delay, $v_k(t) = v(t - k\Delta t)$. Taps are spaced at points equivalent in delay to the Nyquist interval, Δt. Therefore the output at each tap at time t corresponds to a specific point of the image:

$$v_k(t) = v(t - k\Delta t) = r\left[x - (k \bmod n)\,\Delta x,\ y - \frac{(k - k \bmod n)}{n}\,\Delta y\right]. \qquad (4\text{-}10)$$

The cross-correlation with a weighting function is performed by setting the gain at each tap to the desired value $w_k = w(-k\Delta t)$, and summing

$$f(t) = \sum_k w_k\, v_k(t) \qquad (4\text{-}11)$$

$$= \sum_k w(k\Delta t)\, r\,(t - k\Delta t) \qquad (4\text{-}12)$$

$$= \sum_k w\left[(k \bmod n)\,\Delta x,\ \frac{(k - k \bmod n)}{n}\,\Delta y\right] \cdot r\left[x - (k \bmod n)\,\Delta x,\right.$$

$$\left. y - \frac{(k - k \bmod n)}{n}\,\Delta v\right]. \qquad (4\text{-}13)$$

Let

$$i = k \bmod n \qquad (4\text{-}14)$$

$$j = \frac{(k - k \bmod n)}{n} \qquad (4\text{-}15)$$

$$\therefore f(t) = f(x, y) = \sum_i \sum_j r\left[x - i\Delta x,\ y - j\Delta y\right] \cdot w(i\Delta x, j\Delta y). \qquad (4\text{-}16)$$

Various delay line cross-correlator devices have been developed for radar and other signal processing applications and may be applicable to the target recognition problem. Two particular types of devices have special merit. A detailed discussion of optically-tapped acoustic delay line cross-correlators is given further on in this section. These have the merit of being able to handle large numbers of weighting functions in parallel, with no loading problems. Also, weighting functions can be changed readily by changing photographic film mosaic weighting masks. For cases where the

retinal element data is preprocessed into a binary function such as the edge or no edge function previously described, shift register cross-correlators may be used effectively. The principal merit of the shift register cross-correlator is its high state of development for such applications, e.g., the Philco-Post Office Mail Sorter, plus its ability to stop on command and hold an entire array of input data in storage. This latter property is particularly useful in laboratory research studies where the precise nature of the input signal at time of recognition often must be determined.

Summary of Evaluation of Scanning Devices

The material which follows summarizes the conclusions of a survey on methods of scanning photographic data. The survey was conducted to determine the feasibility of scanning aerial photographs having 5000 TV line resolution across a 9-inch field. The video bandwidths involved are commensurate with 9 by 9-inch frames at frame rates on the order of a few seconds. The details of the survey are presented in Appendix E.

Table 4-1, "Comparison of Scanning Devices," lists the candidate scanning devices and evaluation of each in terms of the characteristics considered to be significant in the present application. The devices are commented upon in detail in the paragraphs which follow.

The 5000-element requirement means that six image orthicons would have to be used in parallel, or provisions made for mechanically repositioning the film or scanner in order to cover the full frame. Either case would be undesirable, the first because of image orthicon cost, and the second because of the mechanical problems involved.

The vidicon has about the same resolution as the image orthicon; however, vidicons are relatively inexpensive, so the duplication would not be so costly. The vidicon is slow; 0.1 second or more is required between successive scans, and the viewed object must remain essentially motionless during this time. In addition, there is no possibility of overlap scanning in a single vidicon.

The non-storage property of a flying spot scanner permits it to scan a moving frame. In addition, scanning pattern flexibility and high resolution are available, and may be utilized to advantage in imagery scanning. If, for example, a rotating phosphor scanning tube is used, it is possible to scan horizontally for 5000 elements without the need for duplication of scanning equipment. The element rate of a flying spot scanner is limited by the phosphor decay time; however, fast decay phosphors such as P-16 allow data rates as high as 32×10^6 elements per second.

TABLE 4-1

COMPARISON OF SCANNING DEVICES

		Image Orthicon	Vidicon	Flying Spot Scanner	Image Dissector
Resolution TV lines in screen diameter*		900 TV lines (2%) present 1500 TV lines (50%) development	1200 TV lines (2% response)	9000 TV lines (5% response)	3000 TV lines (25% response)
Speed Capability	elements/sec	50×10^6 (est)	15×10^6	32×10^6	10×10^6
	frames/sec	50 (est)	10	high	50 (est)
Signal-to-Noise Ratio		39 db (4.5 megacycles)	40 db (5 megacycles)	40 db (16 megacycles) 60 db (5 megacycles)	2 to 1
Scan Pattern Flexibility		Inflexible due to storage	Inflexible due to storage	Flexible	Flexible
Cost		Medium	Low	Medium	Medium

*NOTE: Resolutions are quoted at the specific percentage responses for which data was available. Percentage response refers to the ratio of peak-to-peak light output on a test chart at the stated resolution to that at resolutions approaching zero.

A high-brightness, high-resolution flying spot scanner could make use of mechanical scanning in the vertical direction. Film could pass through the scanning plane at a constant rate while being scanned horizontally with a ribbon scan. It appears practicable to think in terms of a rate of motion such that 50 percent overlap is obtained using a 60-element high ribbon scan.

The primary difficulty with flying spot scanners is that of obtaining a trace bright enough to provide a good signal-to-noise-ratio from the photomultiplier. The rotating phosphor tube answers this need, but at considerable expense. In addition, these devices are designed for single line scanning and may not retain high resolution when wide ribbon scan is used.

The image dissector is capable of up to 3000 lines resolution. However, it cannot provide high resolution, high data rate, and good signal-to-noise ratio simultaneously. For the proposed application, satisfactory resolution and data rate result in an extremely poor calculated signal-to-noise-ratio of 2 to 1. To improve this, itwould be necessary to increase picture element size in proportion to the signal-to-noise ratio increase required. A flying spot scanner is the best available type of electronic scanning device. The other three devices are undesirable for the following reasons:

Vidicon -- lag time, resolution, and storage property

Image Orthicon -- resolution and storage property

Image Dissection -- signal-to-noise-ratio and data rate

The remaining question is whether or not a conventional flying spot scanner may be constructed utilizing available high resolution cathode ray tubes, or whether specialized equipment must be developed (perhaps utilizing the rotating phosphor tube). The results of Appendix F indicate that, through careful design, a conventional flying spot scanner may be constructed that provides adequate signal-to-noise-ratio for Laplacian and gradient detection. Therefore, a carefully designed conventional flying spot scanner is recommended for electronic scanning.

Cross-Correlation With Sequential Accessing

A recognition system which scans the input image and inserts it sequentially bit by bit into a dynamic storage device and correlates the contents against a set of reference patterns, has the advantage of effectively testing the correlation between the input data and the templates in all possible translational registrations.

The storage element can be a shift register, delay line, or device which converts the serial data into a parallel accessed array; the required capacity is 1000 bits, at a data rate of 10 megabits per second. To perform the correlation the storage element must be connected to a semipermanent memory containing reference templates.

Shift registers are available in solid state, thin film, and ferrite core configurations. Core register shift rates are less than a megacycle and therefore, inadequate for this application. The thin film shift register has higher speed and high density. Available units shift at a 2 megacycle rate and have capacities of 256 bits. Future development of this device may increase the data rate sufficiently to qualify it for this application. Solid state registers can be obtained with shift rates up to 20 megacycles and capacities up to 300 bits. These registers are relatively expensive and the problems in increasing the capacity to 1000 bits are apparently great enough that some suppliers have refused to attempt more than a 300-bit size. A 1000-bit, 10-mc register would need to be developed and would cost at least $20 to $50K.

A Golay lumped-constant delay line has been built with a 2 megacycle bandwidth, and 200 taps. The attenuation of the line was about 35 db but indications are that this could be reduced. A 1000-tap line with 10 megacycles bandwidth would be a considerable advance in the state of the art; the parts alone will cost in the $10 to $20K range. Magnetostrictive delay lines are unattractive because of the high attenuation (3 db) per tap and the mechanical and electrical problems connected with multiple read-out. Glass delay lines using optical read-out are available with 3 to 5 megacycle video bandwidths and in delays of up to 200 μsec as a practical limit per unit. The attenuation over the length of a 100 μsec line is only 3 db, (the principal losses being in the input transducer). To the cost of the delay line itself, of course, must be added the cost of auxiliary optical and electronic equipment for driving and reading out.

The library of stored patterns against which the input data is to be correlated must be parallel accessible. Photographic film offers high density, analog or digital data storage, low cost, is easily changed, and is read-out with light. Resistive matrices have been used in recognition apparatus, but they are rather expensive and tedious to construct; a rough estimate of cost would be $5 to $10K. Diode matrices fall into a similar category except that the storage is restricted to binary information. Capacitor card, metal card magnetic, and ferromagnetic semipermanent stores have been described in the literature. They are still principally developmental and therefore not readily available.

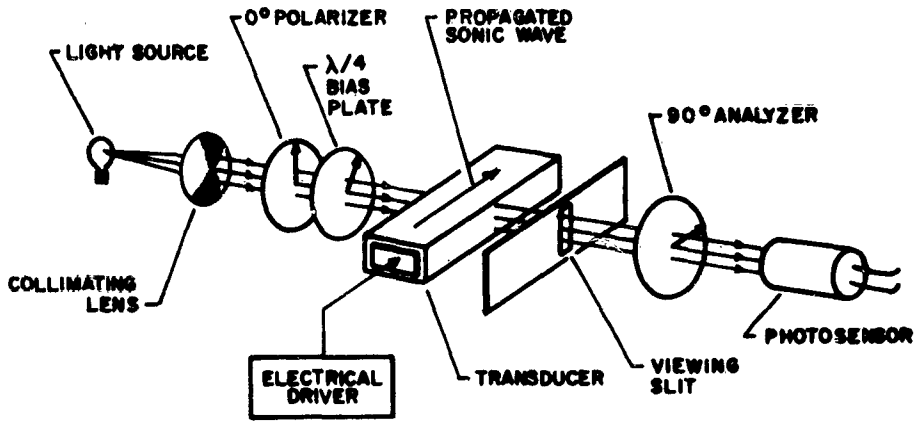
Because the photoelastic glass delay is capable of high data rates, high storage capacity, and parallel read-out at reasonable cost, it appears to be a most attractive candidate for serious implementation study. Current state-of-the-art affords data rates up to 10 megabits per second and a total storage of 1000 bits for a 100 μsec line. The fused silica line is approximately 15 inches in length and should cost from $1,000 to $2,000. This does not, of course, include the necessary optics and driving electronics.

Previous application of the photoelastic delay line has been for the purpose of obtaining a delay device which can be readily tapped either singly or multiply at any desired position on the line. The use of the device described here is different; the operation expected is similar to that of a large capacity shift register with all bit positions available simultaneously for parallel output. This type of operation will require experimental work to ascertain its practicability and to assess the magnitude of the problems involved.

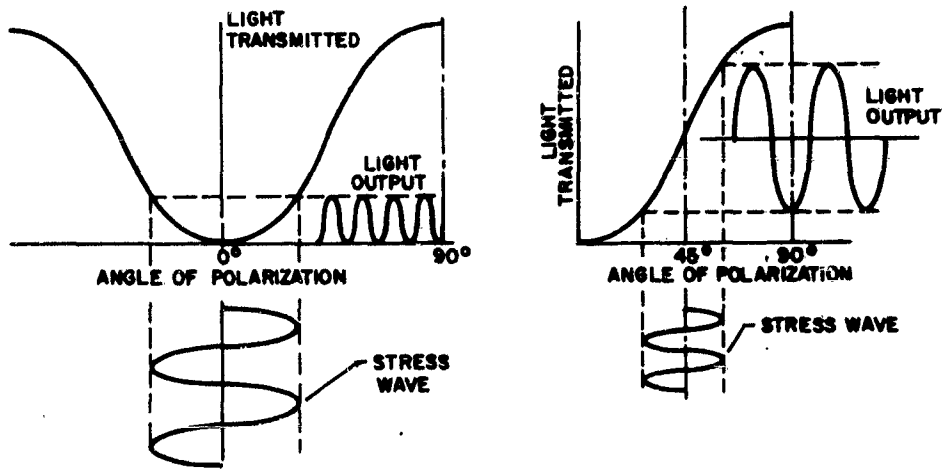
The fact that the glass delay line seems to afford the best solution to the requirement for a multiply tapped, large-capacity storage device, coupled with the fact that photographic film is the most flexible and economical medium for the semipermanent template store, leads quite clearly to the highly compatible combination of the two as a preferred implementation. The electrical interconnection of 200,000 points in the template store and shift register is replaced with an optical interconnection; the glass delay line not only provides the shifting and data storing function, but provides the light source directly for interrogation of the photographic memory bank.

Figure 4-6 shows a typical arrangement for optically tapping a glass acoustic delay line. A ceramic transducer bonded to one end face of the glass slab propagates an acoustic wave through the length of the bar. An absorber is bonded to the other end to prevent reflection of the wave after it has traversed the length of the bar. An acoustic signal is propagated as a mechanical stress; therefore, the wave traveling down the glass is a moving stress pattern. Glass, and certain other materials, possess the property of becoming optically birefringent when subjected to mechanical stress; that is, the glass exhibits different indices of refraction for light polarized parallel to or perpendicular to the direction of stress. This produces a change in phase relationship between two mutually perpendicular polarization components of light projected through the glass.

Referring to Figure 4-6, a light source and lens produce a collimated beam of light, which is projected through the glass bar in a direction perpendicular to the propagated sonic wave. A viewing slit is placed on the far side



Basic Elements for Optically Tapped Delay Line (Single Bit Readout)



Operation Without Bias Plate Operation With λ/4 Bias Plate

Figure 4-6 Optically Tapped Delay Line

to mask out all the light except a narrow aperture which is adjusted in width to be a fraction of a wavelength of the RF carrier in the glass medium. A polarizer is inserted at the input face of the line and an analyzer is placed at the output face with its orientation adjusted to completely cut off the transmitted light. With the crossed polariods, no light will be seen by a photomultiplier which views the slit as long as no stress is introduced into the glass bar. Acoustic stress waves traveling down the line, however, will effectively rotate the polarization of the light as it passes through the bar and produce a light output. The transfer characteristic is indicated in Figure 4-6.

In order to obtain a true reproduction of the sonic wave in the glass bar, a quarter wave bias plate is introduced to shift the operating point to the center of the characteristic as indicated in the second transfer curve in Figure 4-6. Best linearity and sensitivity are also obtained at this point. The delay can be easily adjusted to any value desired by moving the bar with respect to the rest of the optical components and thus changing the point at which read-out occurs.

A diagram of the general optical correlator system is shown in Figure 4-7. It is assumed that a scanner and preprocessor generate a continuous ribbon of input data from the scanned photograph and that the input to the correlator is an array of bits describing each picture element by means of a ONE or a ZERO (it could as easily be analog, however).

The input data is used to modulate a carrier which is propagated as an acoustical stress wave down the glass delay line by means of the ceramic or quartz transducer cemented to the end. A collimated polarized light source of sufficient size to illuminate the entire length of the bar projects light through the glass at right angles to the direction of wave travel. When viewed through an analyzer and a slit at each bit location, the entire contents of the delay line appear as a moving array of modulated light patterns. An optical system can project this data array onto a bank of photographic mask sets. A better method, which eliminates the optics, is shown in Figure 4-7 and places the photographic mask sets directly after the analyzer and slit array on the output face of the delay lines.

The modulated light coming from each bit position in the input data array is transmitted through the photographic masks for all recognition classes; only two classes are shown in Figure 4-7 for simplicity. Each class mask is divided into a specific area for indicating positive weighting for any

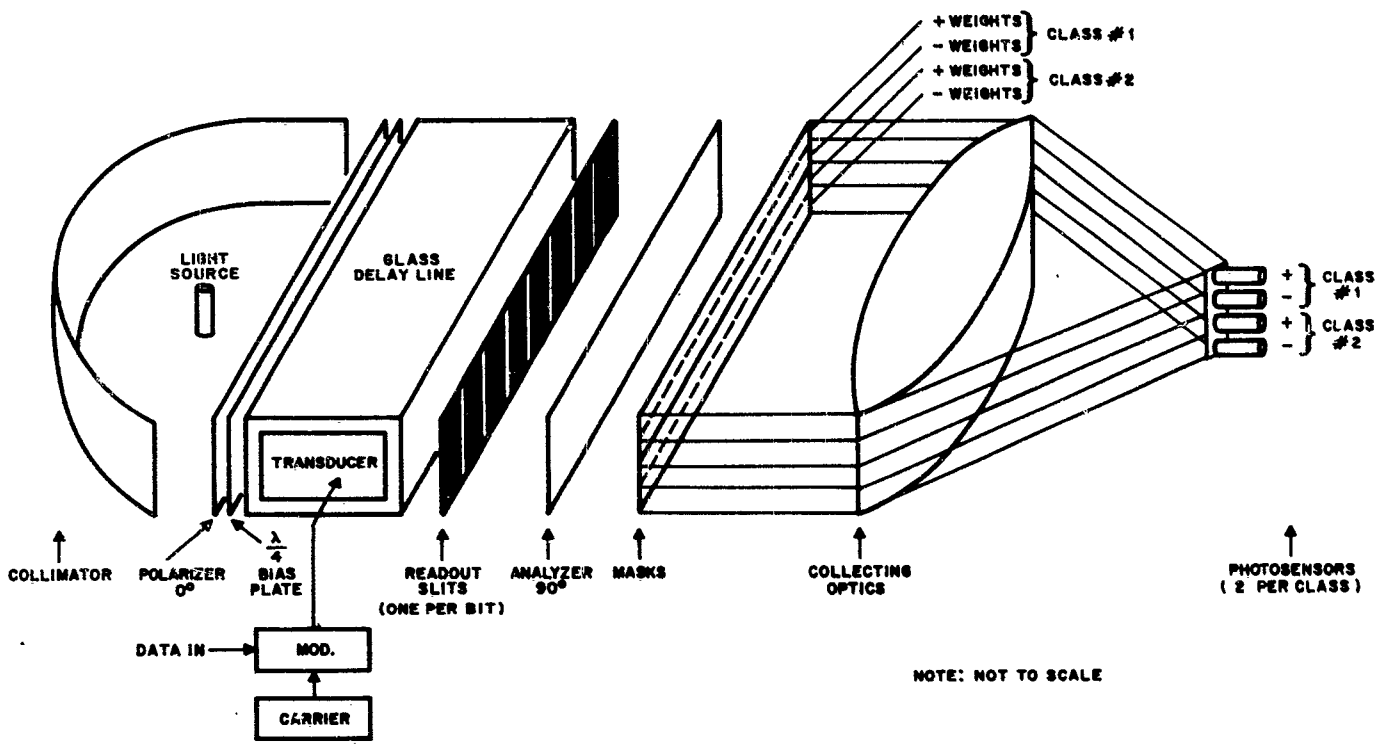


Figure 4-7 Optical Correlator With Two Sensor Summation

bit location and an area for negative weighting. The magnitude of the weight to be applied is determined by the degree of exposure of that particular bit location on the mask during the preparation of the mask. The correlation of each bit in the input data array with the stored reference pattern on the mask is therefore simply the amount of light transmitted through the mask. The sign, positive or negative, is indicated by the location of the weight in either of the two specified mask areas; if the weight is positive, the negative weight space is black, and the positive weight space is clear and vice versa when the weight is negative. Correlation of the input pattern with the stored pattern is taken as the algebraic summation of all the positive and negative bit correlation quantities. Two photosensors, with their necessary optics, are arranged to view the positive weight areas and the negative weight areas respectively, for each class, over the entire length of the delay line. This is shown schematically in Figure 4-7 as a cylindrical lens system designed to reduce the length of the viewed area to a dimension compatible with the photocathode size in the viewing photosensor. It is anticipated that a fiber optic assembly may be superior to a lens for collecting the total transmitted light and performing the appropriate change in geometry. This is discussed in Appendix G. The photosensor outputs for each class are then subtracted electrically and a threshold applied for making an output decision.

The advantages of a system of the type under discussion here lies in the possibility of simultaneously correlating a large section of the serially presented input data with a very large number of reference patterns. The number of such classes is limited only by the practical problems of registry, class separation, and physical location of the required photosensors. Obtainable data rates also appear adequate for the desired recognition rates.

It is highly desirable from the standpoint of threshold stabilities to be able to sum the positive and negative mask correlations algebraically without the difficulties usually experienced in subtraction circuitry balance and symmetry. Several methods have been proposed and examined for avoiding the drift and device matching problems attendant to photocell bridge and electrical differencing circuits. At least two of these appear sufficiently practical to warrant serious attention; they are a carrier commutated light subtraction technique and an interdigitated grating light commutation arrangement. These methods are described in detail in the following paragraphs.

Figure 4-8, "Single Sensor Optical Correlator Using Carrier Commutation," shows a system which is different in only two respects from the original two-sensor arrangement. First, only one photosensor, viewing both the positive and negative mask areas, is used. Second the analyzer has been replaced with a composite which equips all positive weight areas with a

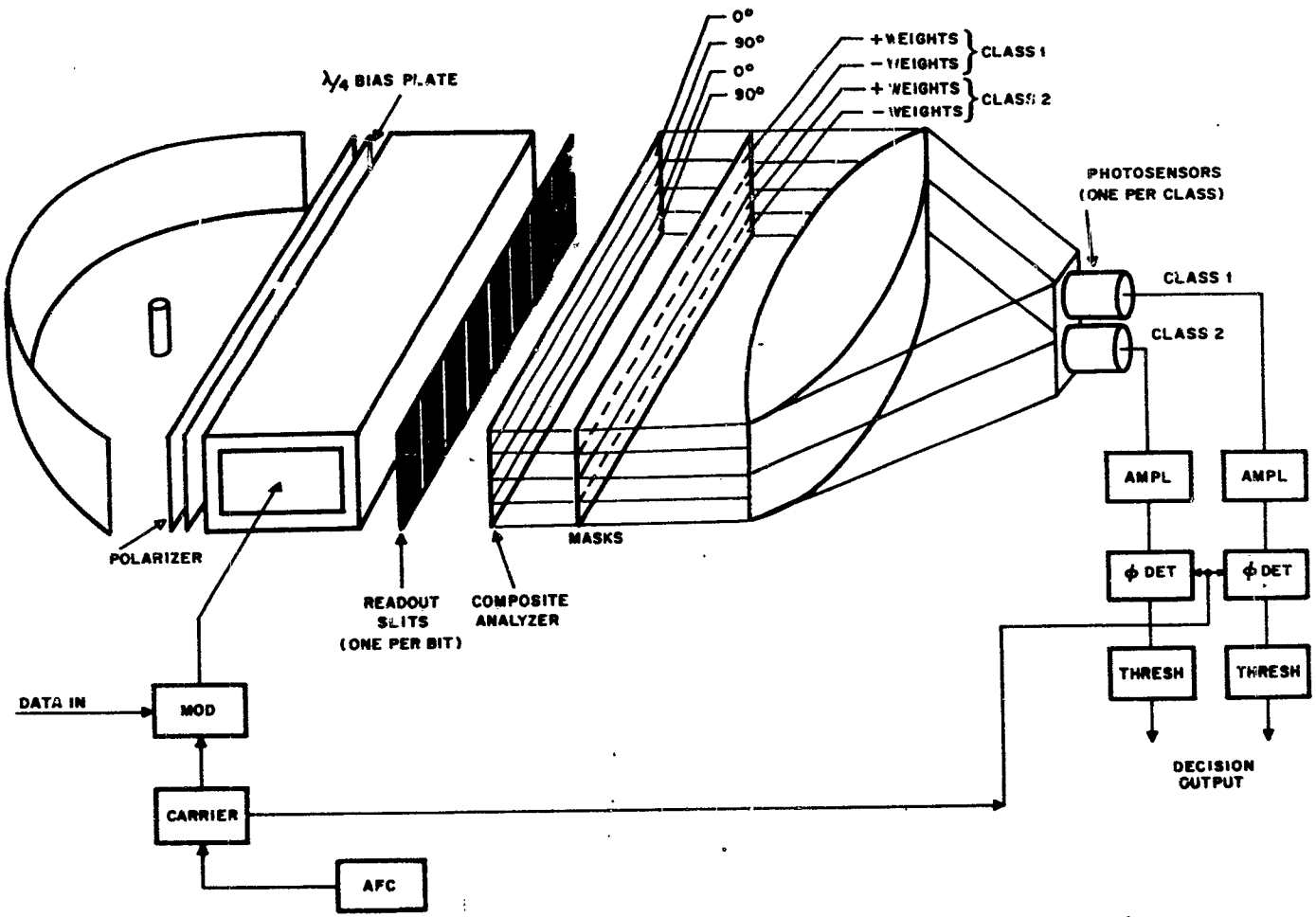


Figure 4-8 Single Sensor Optical Correlator Using Carrier Commutation

polariod analyzer at the same orientation as the original, and equips all negative weight areas with an analyzer oriented at right angles to that of the original.

To explain the effect of these changes, consider the basic single slit read-out arrangement described earlier. The effect of the stress variations traveling past the read-out slit was to perturb the vector, representing the light polarization about the 45° bias point. When viewed with the analyzer, this gives the 0° transfer characteristic shown in Figure 4-8a. If the analyzer in this original arrangement is rotated through 90°, the transfer characteristic becomes that shown for 90° in Figure 4-8a. The light intensity variation at the carrier frequency as seen by the photomultiplier is now of reversed polarity for the 90° analyzer position compared to the original 0° position.

Another way of looking at it is illustrated in Figure 4-8b. Assume the mean position of the output light polarization to be at 45°. Two analyzers illuminated by this light and arranged at 0° and 90° will analyze the incident light into its 0° and 90° components. Movement of this input light vector caused by the stress wave in the glass causes the variation in light transmitted by the two analyzers to be of opposite phase, i.e., when the light through the 0° analyzer is instantaneously increasing, the light through the 90° analyzer is decreasing. A photosensor which views the light variations transmitted by both analyzers observes the difference. The polarity and magnitude of the photosensor signal can therefore be determined by introducing a mask between the two analyzers and the sensor in order to alter the relative transmission of the light paths. Only one sign of weighting is used at any one-bit location; the other sign has an opaque mask. A full positive or negative weight would use a mask which was transparent in one analyzer area and opaque in the other. The light shown coming through both masks in Figure 4-8b actually refers to the summation of light from all the masks of each sign.

The carrier-commutated system requires good registration of the viewing slits with respect to the carrier wave in the delay line, in order that the light from all slits will add in phase for each direction of polarization. Since an array of 1000 such slits occuping 15 inches is contemplated, a high degree of accuracy is required. The higher the carrier frequency, the greater the required slit location accuracy. It may be necessary to provide the system shown in Figure 4-9 with an automatic frequency control loop to maintain the carrier in registry with the slit array.

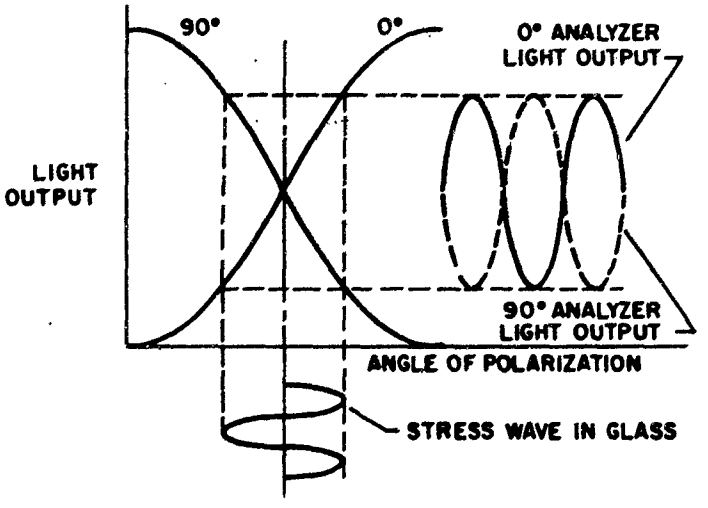


Figure 4-8(a) Transfer Curve for Split Analyzer

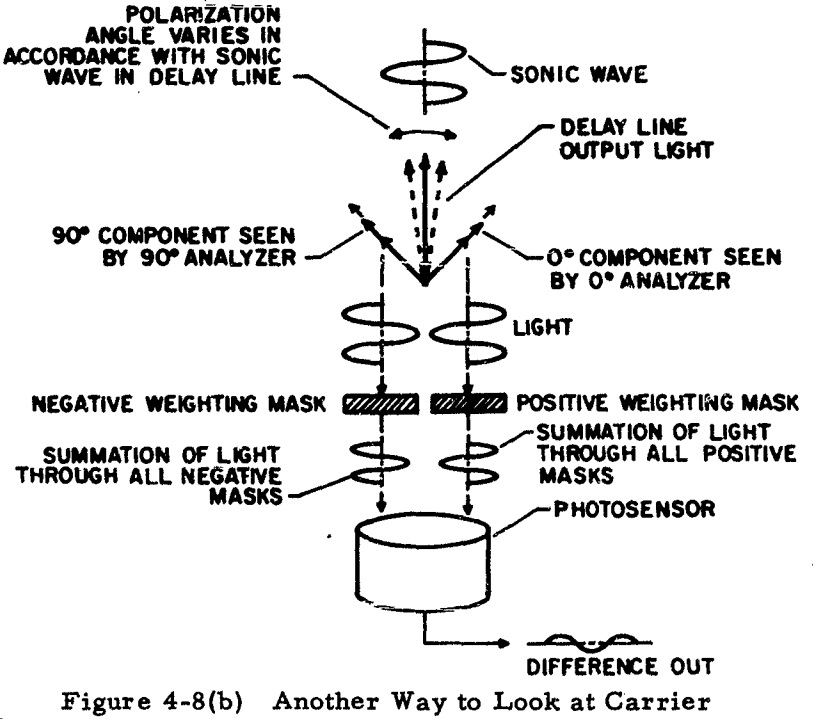


Figure 4-8(b) Another Way to Look at Carrier Commutation Subtraction

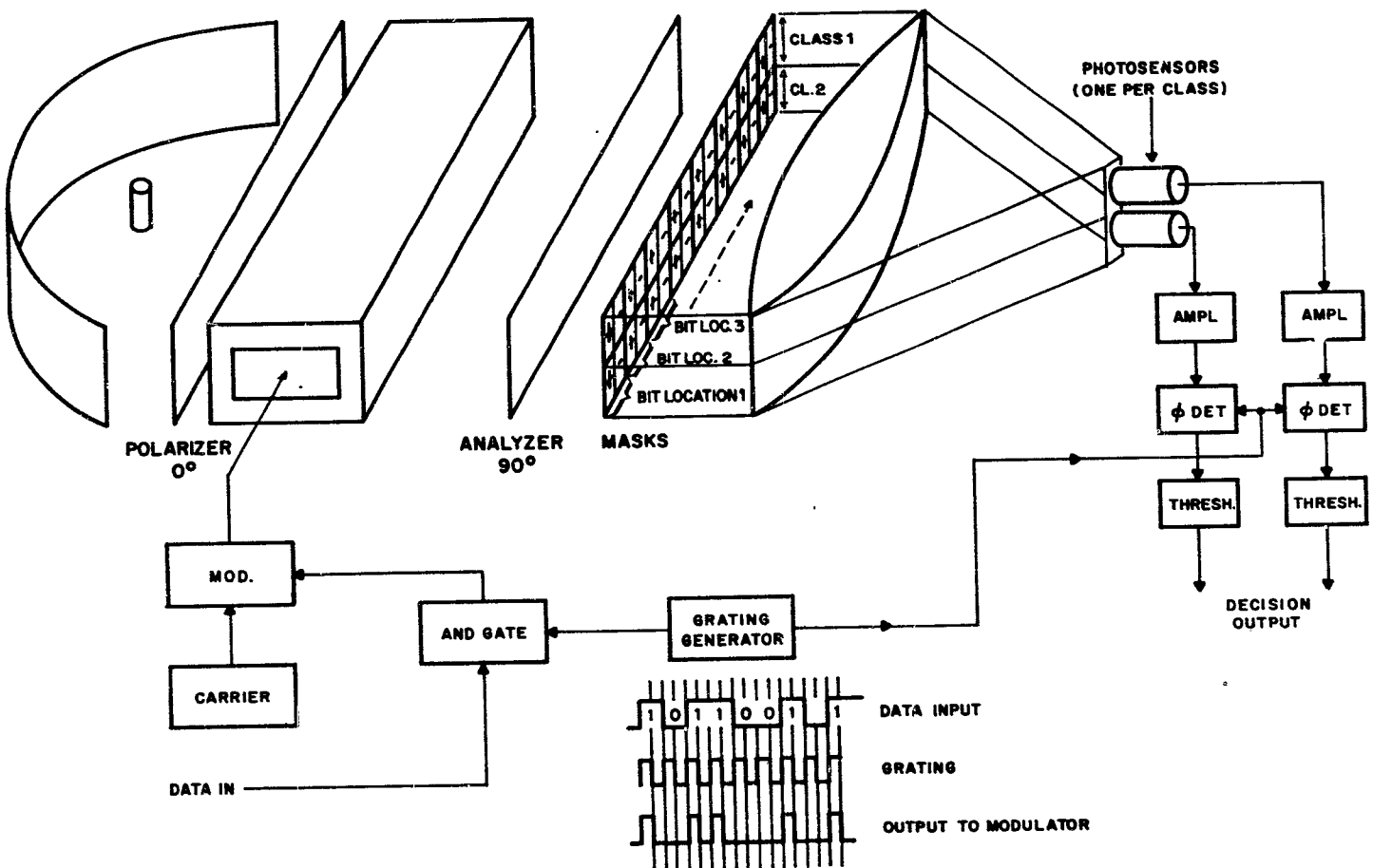


Figure 4-9 Single Sensor Optical Correlator Using Interdigitated Commutation

A simpler method appears feasible for providing single sensor algebraic summation of the positive and negative correlations. It is based on the use of a mechanism for alternately presenting the positive and negative mask correlation quantities to the photosensor; an illustration of this is shown in Figure 4-9. The glass delay line is uniformly illuminated by a collimated polarized light source, as before, and the data are again used to modulate a carrier which drives the transducer. The quarter-wave bias plate is omitted, however, to place the operating point at minimum transmission. A burst of carrier, therefore, causes a full-wave rectified light output (see Figure 4-10) when viewed through the analyzer.

In order to produce alternate presentation of the positive and negatively weighted light quantities to the photosensor, a moving grating is interleaved with the input data. Each data bit is reduced to half its former duration and a ZERO (no light) is inserted between each bit. The signal driving the line is shown in Figure 4-11.

Every other bit location in the input signal is made a ZERO to form the grating; the remaining bits represent the data by inserting an RF burst for a ONE and no RF for a ZERO.

The mask format contains two possible positions for weighting at each bit location. The bit location is divided in half. For positive weighting, the first half is used; for negative weighting, the second half is used. In other words, the mask for each class weights each bit either positively or negatively by blacking out one of the two halves at each bit location and inserting the weight value in the other half by shading.

The effect of the moving grating and the stationary mask format is to first transmit to the photosensor the input data multiplied by all the positive mask weights, and then by all the negative weights.

The alternating component of the photosensor output is then phase detected with the signal used to form the grating in order to determine the sign and magnitude of the total algebraic summation for each class.

Interdigitated grating operation is much simpler than carrier commutation. No slit array is required, which avoids fabrication and registry problems, and binary light operation is used which permits some improvement in bit resolution (see Appendix G). This is obtained at the expense of using only half the bit capacity of the delay line for holding the data. If it is assumed that the same data input is used for this arrangement as for carrier

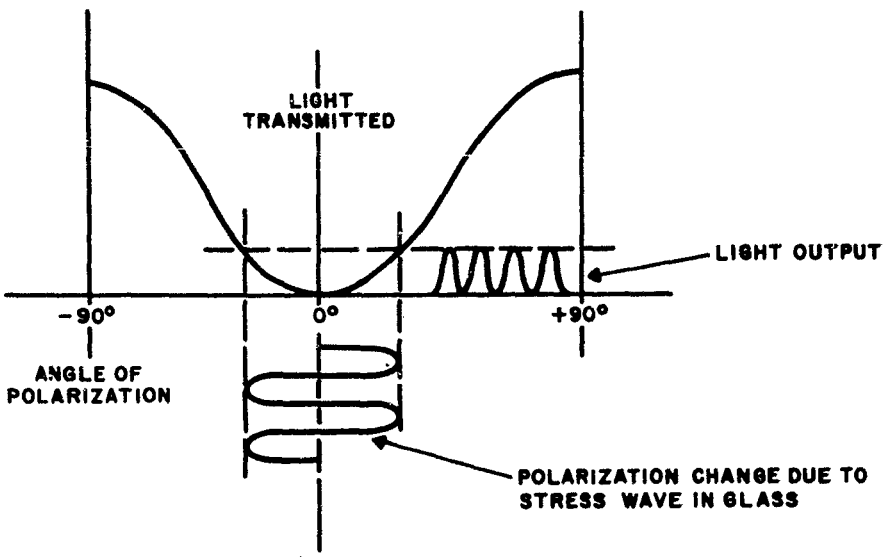


Figure 4-10 Transfer Characteristic Used for Interdigitated Grating Commutation

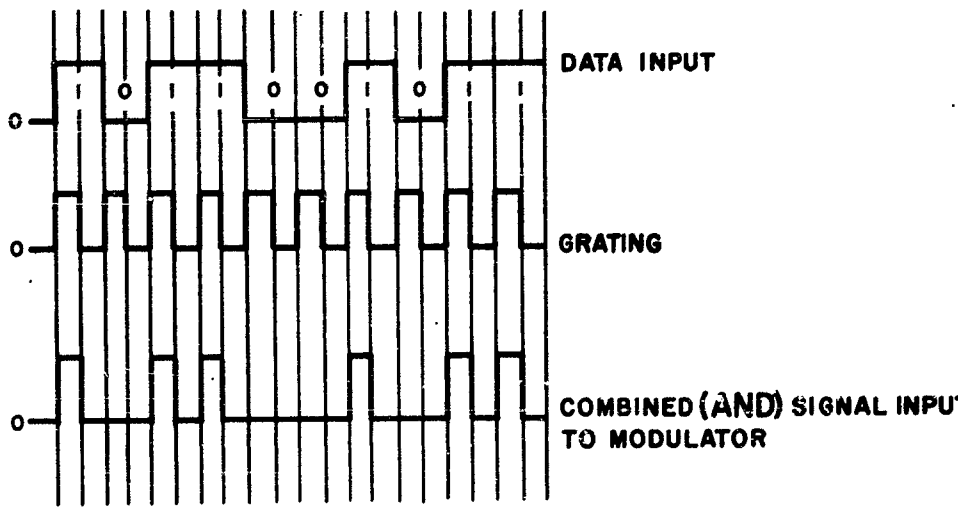


Figure 4-11 Modulation Input to Delay Line

commutation, the effective bit rate necessary in the delay line for the same data rate becomes exactly doubled. If the data input contained all ONES, for example, the delay line contents would be a train of alternating light and dark stripes at twice the bit rate. This approach is attractive for high frequency transducer operation. For a 10-megabit data rate, a carrier frequency of 40 mcs and a transducer bandwidth of 20 mcs would be adequate. Since slit read-out is not used, there is no problem of phase variation at high carrier frequencies because of slit misalignment, as there may be with a carrier commutated system.

SECTION 5

CONCEPTUAL DESIGN OF AN OBJECT RECOGNITION SYSTEM

5.1 General

Figure 5-1 presents a conceptual block diagram of a system tentatively proposed for the image screening application. Designing and evaluating such a system provides a test of the validity of the concepts deriving from the study. The fundamental building blocks of the system to be considered here are the following:

- a high definition flying spot scanner with digitally controlled film transport and sweeps
- electronic preprocessing and recognition circuitry
- digital logic to combine outputs of recognition sub-units, code the outputs properly, and command the scanner
- a means for recording coordinates of detected objects
- an off-line display for convenience of the photo-interpreter
- a control console

The discussion which follows treats each of these items in some detail.

5.2 Flying Spot Scanner

Since high-quality aerial photography may involve as many as 5000 separate resolution elements across a 9-inch dimension, a high resolution sensor is required. A summary of the choice of scanning sensors available and the many considerations involved in the choice of a sensor is presented in Appendix E. At present, the best sensor choice appears to be a twin flying-spot scanner as shown in Figure 5-2 working with the semitransparent film negative. The use of a twin scanner is necessary because flying spot scanners with overall resolutions of 5000 television lines or better lie near the frontier of the art. When 5000 lines are obtained they are achieved at the cost of low beam current, thin phosphor, and low brightness output.

In the proposed twin scanner, each section has an overall resolution in excess of 3000 lines, in order to achieve a total resolution of 5000 lines or better. By working directly with negatives or positive transparencies, the

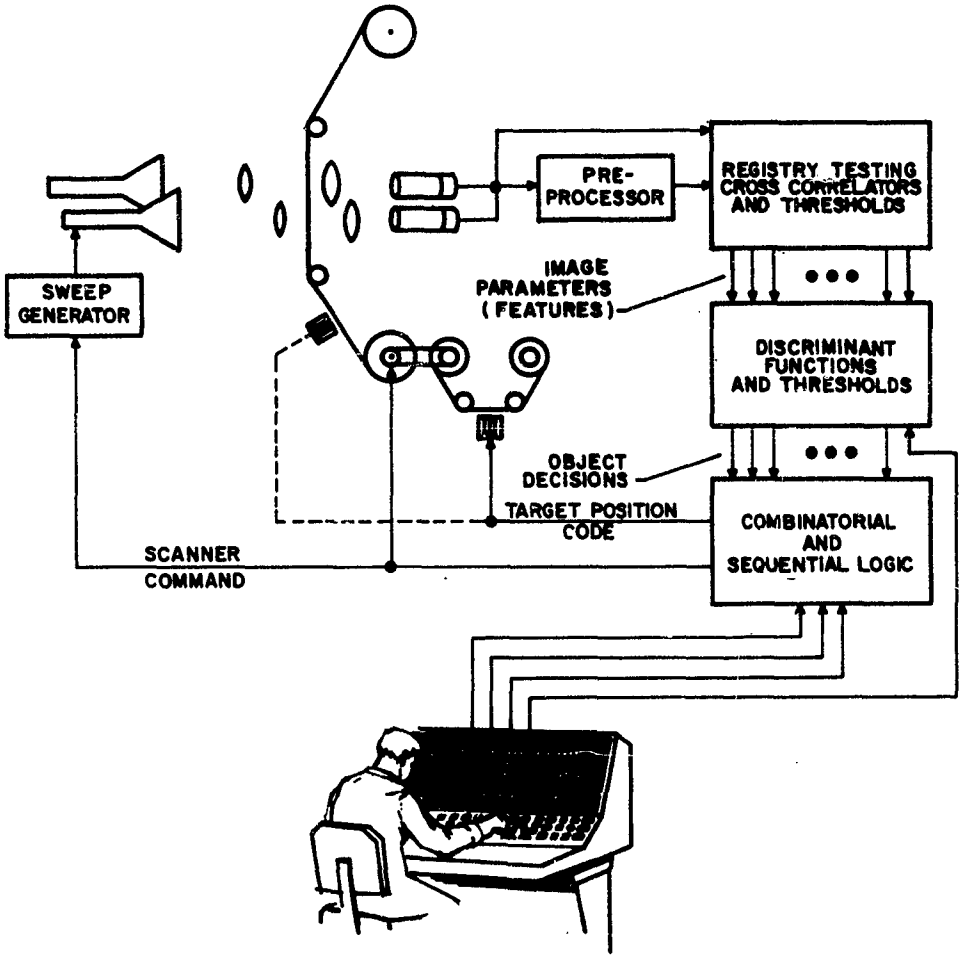


Figure 5-1 Conceptual Block Diagram of Image Screening System

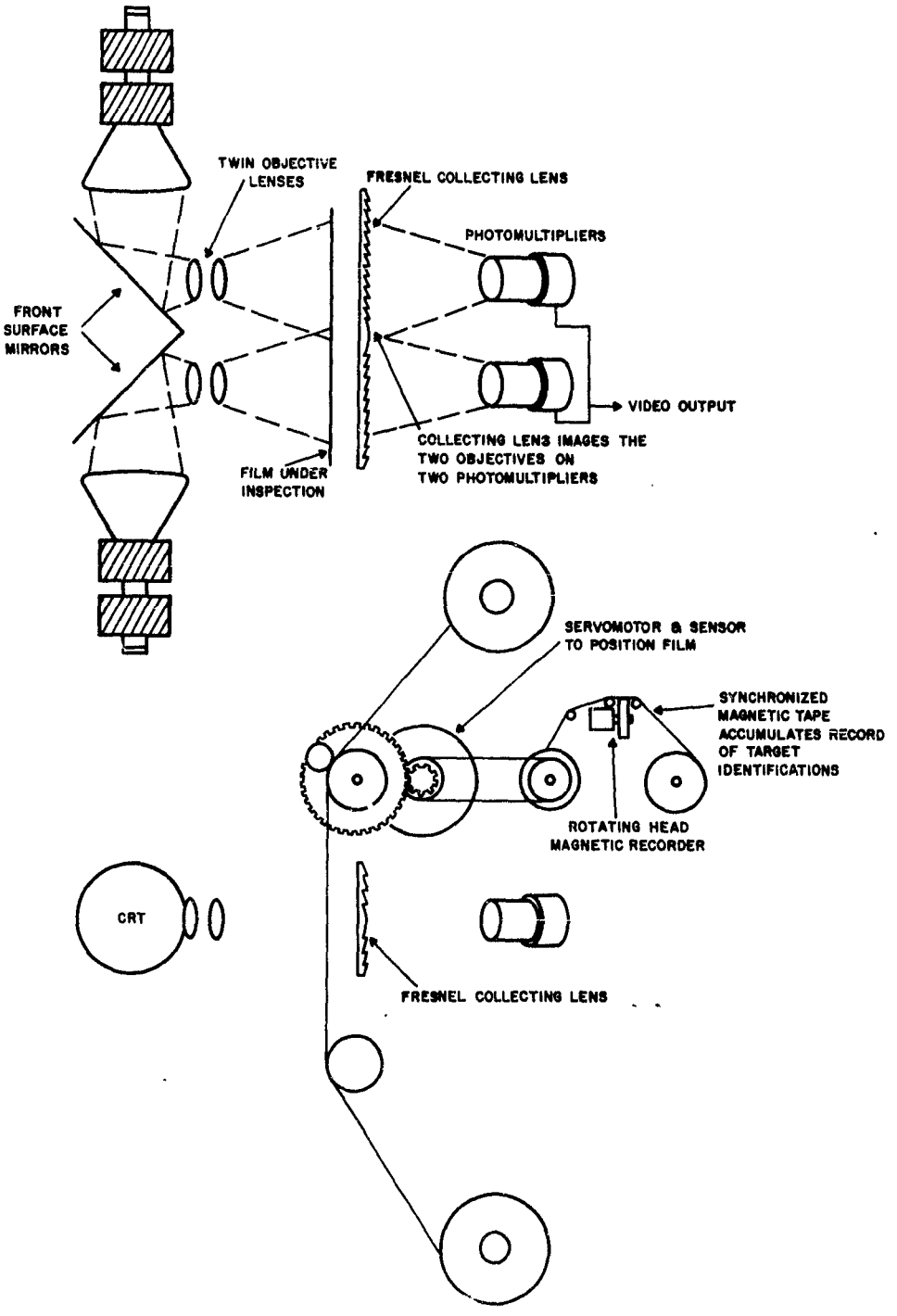


Figure 5-2 Scanner and Film Transport

system obtains light collection efficiencies as high as 100:1 greater than could be obtained with opaque prints. Appendix F shows that by using P-16 phosphor in the CRT, a video bandwidth of five megabits should be possible with a signal-to-noise ratio of 40 db. This data rate corresponds to one picture element every 0.1 μsec, or an entire 9 by 9-inch frame every 2.5 seconds, if every picture element is scanned once.

The film transport can be a digitally controlled servo system to advance the film, and start, stop and reverse at command from the recognition logic. A compensating vertical deflection signal can be obtained from this servo and used on the scanner to compensate for instantaneous positioning errors. The film can either advance continuously or be stopped occasionally for repeated inspection of suspect areas at different scale factors.

The scanner sweeps are periodic, but of variable size and pitch determined by digital command from the control console and recognition logic. Figure 5-3 shows the general form of the scan, which is a sawtooth ribbon progressing across the width of the 9-inch film. Normally, the ribbon proceeds uninterruptedly across the film, crossing over first one CRT and then the other with sufficient overlap in the fields of the two CRT's to ensure recognizing objects which lie in the region of overlap.

The normal size of the scanning ribbon is determined by the maximum dimension of the objects to be recognized as individual entities in the scaled photograph. For example, if the device is programmed to locate armored vehicles, the ribbon will have a vertical dimension twice that of the largest tank diagonal expected in the photograph.

The output of the scanner is 5 mc gray-scale video. This signal is then preprocessed to obtain a quantized signal which emphasizes edges, corners, and straight lines. The presently preferred method of extraction is to obtain a signal roughly proportional to the magnitude of the gradient vector of the brightness signal in two dimensions, i.e., $|\text{Grad } B(x, y)|$. The use of the circuit shown in Figure 5-4 is an effective way to obtain this signal. Appendix A shows that a gradient signal obtained in this way with a tapped delay line has a signal-to-noise ratio only 3 db less than that of the original gray-scale video from which it is derived. Both the original gray-scale information and the quantized gradient signal then become inputs for registry-testing cross-correlators which follow.

5.3 Recognition Logic and Components

The recognition logic which is described here is designed with the following considerations in mind.

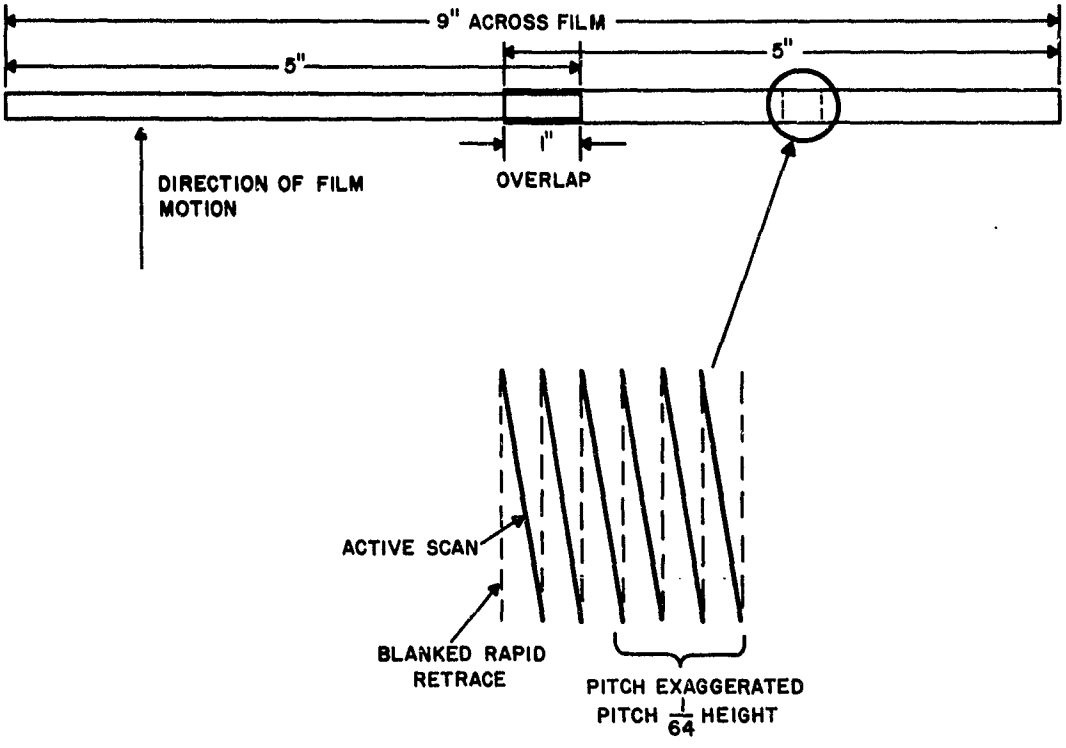


Figure 5-3 Sawtooth Ribbon Scan Used With Twin Scanner

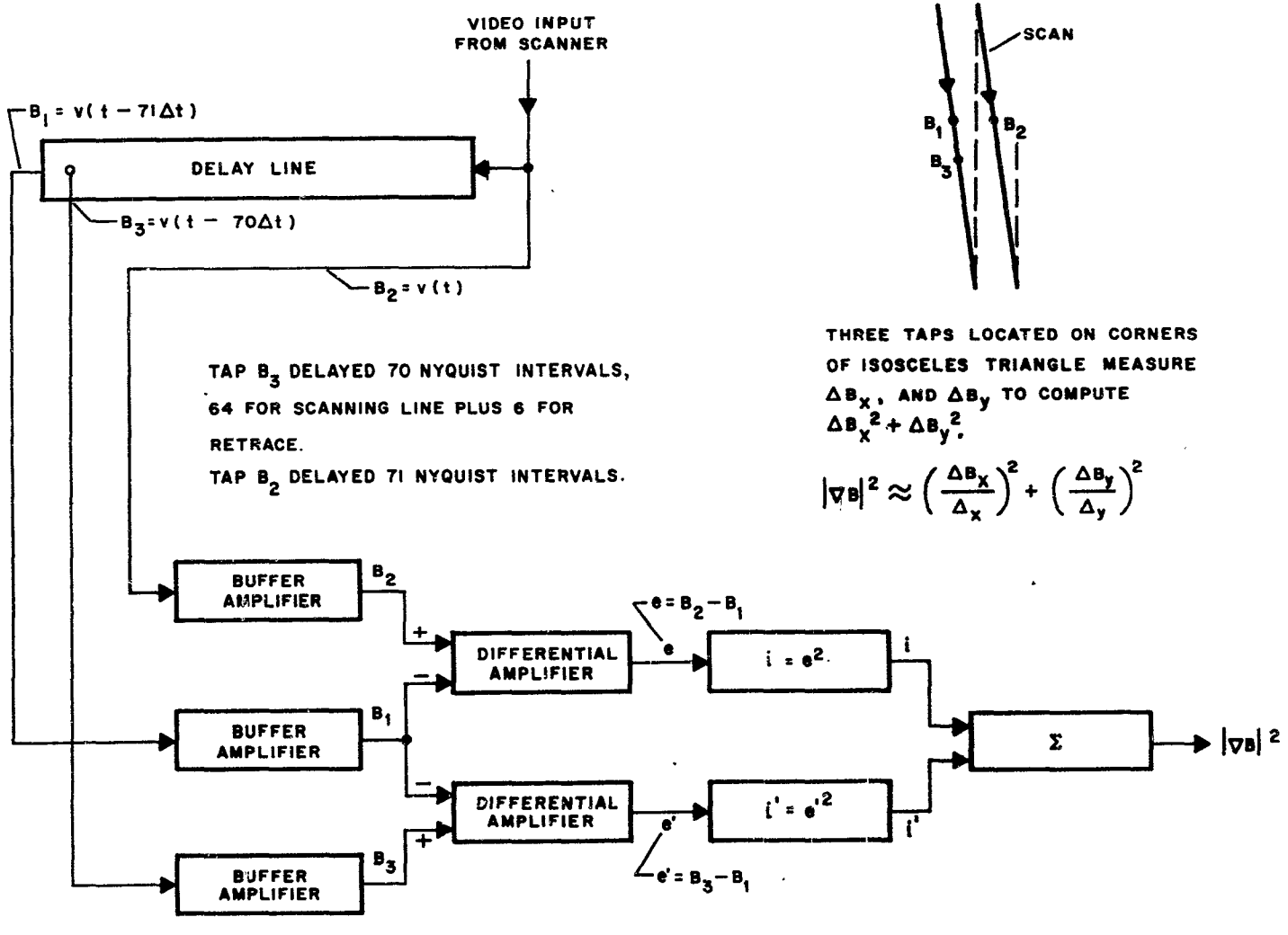


Figure 5-4 Showing Method for Deriving a Signal Proportional to the Magnitude of the Brightness Gradient Vector

- Object recognition circuits and networks are set up to produce a recognition output only when the desired object is centrally located and aligned with the recognition circuitry in optimum translation, with the major axis within $\pm 10^{o}$ of design center for the logic.

- Remaining angles of rotation are covered by duplicating the recognition logic in 8 to 17 other orientations as needed.

- Optimum translation is obtained by permitting the input signals to propagate down a delay line cross-correlator in such a way that input patterns align themselves with every possible position of translational registry as they propagate down the lines. Threshold recognition circuits and all combinational logic which follows must then have sufficient switching speeds to respond with indicated recognition outputs at any of the many positions of registry through which the input image passes.

The organization of the recognition logic is of the form shown in Figure 5-5. An initial layer of threshold logic provides local area feature detection. These local area features together with a few important feature pairs (generated by AND gates) are then combined with linear weights (a linear discriminant function) to achieve a threshold object detection. Eight or more of of these individual object detections, each for a different range of angle, are then combined in a multilegged OR gate to indicate final object recognition. Although only one delay line is shown in the diagram in Figure 5-5, it should be understood that at least two such delay lines will be used, one containing gray-scale information and the other containing quantized gradient, with appropriate feature detectors on each.

The delay line cross-correlator and associated feature detectors are the heart of the object recognition system. It is the ability of this parallel network to perform a large number of linear summations and threshold decisions per unit time which makes the device effective for rapid object recognition. Consider a delay line containing 1000 separate resolvable elements, each capable of being individually weighted and brought to a threshold decision. Perhaps only 100 of these elements will actually receive weights for any one threshold input. Each picture element occupies 0.1 μsec space in the line, and a new summation and threshold decision is achieved every 1/10 μsec. Thus, one threshold element performs an analog summation of 100 variables in 0.1 μsec, for a total of 10^{9} analog variables summed per second. A single object recognition may depend on the parallel operation of 100 of these threshold summers; thus 10^{11} analog variables are summed per second, and subjected to 10^{9} separate subdecisions per second in order to scan the photograph for a single candidate object.

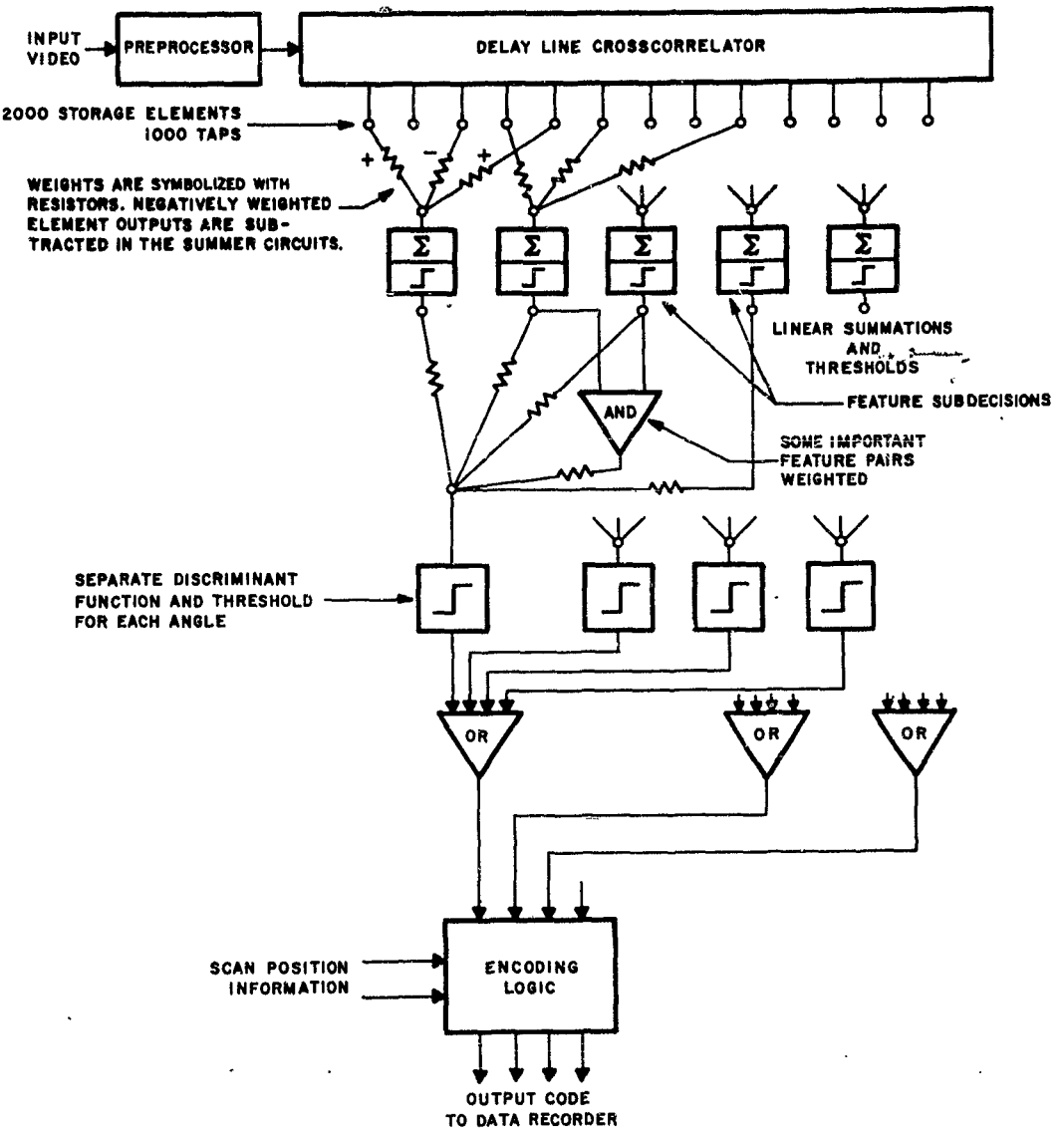


Figure 5-5 Organization of Recognition Logic

Although the object recognition equipment is here described in terms of a delay-line cross-correlator, a shift-register cross-correlator could be used as well. Appendix G describes an optically tapped acoustic delay line as a cross-correlator for object recognition applications. The optically tapped delay line appears at this time to be the best choice for the following reasons:

- The basic cost per buffered storage element is believed to be an order of magnitude less for the acoustic delay line: $5 per bit versus $50 per bit.
- Changing the weights to summing thresholds for different recognition problems involves merely changing photographic slides in the case of the optically tapped delay line. For a lumped constant line or a shift register, however, it involves changing a wiring assembly which may cost $10,000 to $20,000.
- The assembly of fiber optics and associated weights is presently believed to be more easily achieved than a wiring matrix which permits any of 100 or more thresholds to have access to any of 1000 taps.
- Gray-scale information can be handled more readily in the delay line.

The system described thus far will be capable of performing an elementary image screening function, i.e., recognizing and marking elementary objects which can be represented with sufficient accuracy in 32 x 32 element retinal array. The device scans the film continuously and marks such simple objects as tanks, trucks, helicopters, and tents as they are encountered. By merely changing the size of the scanning standards used, any object or object complex for which 32 x 32 element resolution is sufficient for recognition can be so marked.

5.4 Off-Line Display for Photo-Interpreter

In order to utilize the output of the screening device, some sort of viewing mechanism is required for use by the photo-interpreter which will permit extended viewing of any one frame while the image screening apparatus continues processing photographs. One approach is to have individual recognitions marked on a magnetic tape which passes under the recording head synchronized with the processed film as it passes through the screening apparatus. The marking indicates target type and target coordinates in the 9-by 9-inch frame. When the entire roll has been processed on the screening equipment, the roll and its associated magnetic record can be placed on a viewing table for inspection by the photo-interpreter.

In the design considered here, the 9- by 9-inch film is illuminated from the rear by a diffuse uniform light for viewing. At the same time, the prescreened target locations are indicated by brighter illumination in a local area, and, if desired, different colors of local illumination can be used to indicate different target types. Accurately placed local illumination can be achieved by means of a matrix of small light bulbs placed behind the film. A particular bulb in a particular location would be activated when the digital coordinates for a detected target location correspond with that particular bulb.

A relatively small amount of logic built into the viewing table can give the photo-interpreter considerable freedom of selection. For example, a small number of push button controls can be arranged on the viewing console to make the following options available to the PI.

- Advance the film one frame at a time whenever the advance button is depressed.
- Advance the film to the next frame not obscured by cloud cover when the advance button is depressed.
- Advance the film to the next frame containing a detected target when the advance button is depressed.
- Advance the film to the next frame containing a particular target type (e.g., tank, truck, mobile gun emplacement, helicopter, tent).

The following optional modes are possible for concentrating attention in the display upon particular targets.

- Display film with uniform illumination, no colored illumination for target designation.
- Display film with all detected targets illuminated with their appropriate color codes.
- Display film with particular targets or combinations of targets marked (selection to be made by panel controls).

For the photo-interpreter to make the most effective use of the image screening equipment, he must have the option of varying the recognition threshold for various target candidates. If the threshold is set too low, the P. I. may lose important time searching through an excessive number of false alarms; if set too high, too many targets may be missed. Thus, if hunting for a particular target such as an expected armor concentration, he might set the threshold quite high and quickly scan the available photography for recognitions. Failing to find the required target on the first try, he has the option of lowering the threshold and trying again.

One way to achieve the versatility described above is to make individual threshold controls available at the console of the screening equipment. This technique, however, has the disadvantage of tying up the screening equipment while repeated scans are made on the film with different threshold settings. A better approach is to record target identifications together with a digital number which indicates the relative degree of certainty associated with each identification. The threshold for target designation to the P. I. can then be implemented by logic contained within the viewing console. If searching for armor, the P.I. then might set red lights to mark armor and blue lights to mark possible tracks. The threshold controls on the viewing console would be set to a relatively high value for both indicators. The film-advance mechanism would be set to advance rapidly, stopping automatically on all frames containing either of the two chosen recognitions with probabilities exceeding the preset threshold value. In each case, a touch of the advance button by the P. I. would advance the film to the next frame containing a recognition of the required probability. If the number of valid recognitions verified by the P. I. lies within the range of that expected, the P. I. might elect to go on to another roll of photography; if not, he might prefer to reinspect the film with the designation threshold set to respond to those identifications lying below the previous threshold and above a new one. This process could continue until the number and type of false alarms made further scrutiny seem unwarranted.

5.5 Extended Search Capability for Special Situations

A stored-program digital computer can be programmed, given sufficient memory and time, to solve very difficult pattern recognition tasks. Direct use of the computer on real-time pattern recognition tasks is usually not practical because excessive time is required to process all of the picture data in its many possible combinations. However, if the object recognition system described thus far is used to screen the data entering the computer so that only the more interesting and promising areas are considered, times for computation may fall within practical bounds. In general, the computer could be used in two ways.

First, local areas in the photograph, already designated as probable target objects by the high speed recognition device, could be subjected to further scrutiny involving relatively elaborate sequential routines in order to verify recognition and classification. To do this, both gray-scale and gradient data for the candidate area could be read into computer memory for programme scrutiny. Typically, contours may be followed, different thresholds tried to verify presence of an obscure but necessary feature, and even local area Fourier spectrum components measured. Computation time would be relatively short in this process for the following reasons:

- Object has been pre-centered and pre-rotated in field by previous registry testing recognition logic.

- Only those logical tests necessary for verifying the existence of a particular object need usually be tried. If recognition is not verified then other tests for the next most probable identifications can follow.

A second way in which the computer may be used is in the recognition of target complexes. Verified individual object recognitions and their coordinates may be considered jointly in their various combinations in order to recognize object complexes such as missile launching sites, or multiple gun emplacements. In this mode of operation the computer program could consider the various detected objects within a given area as a property list and then hypothesize a candidate target complex. Further verification could then take place with the scanner under the command of the computer. For example, the scanner-computer combination could be used to attempt to find and trace paths or roads joining gun emplacements and possible ammunition or fueling depots. Such sequential techniques are essential when the total complex encompasses a greater number of picture elements (in the resolution required for recognition) than can be held in memory at any one time.

5.6 Control Console and System

The following adjustments can reasonably be expected to be available at the control console:

- Classes of target objects to be included in the search
- Scale factor adjustment for the photograph plus the size of the scanning pattern
- Search time to be spent per frame
- False alarm rate and detection probability by threshold adjustment

To permanently wire all feature and object recognition logic and weights for all the target classes likely to be used in screening, would lead to almost prohibitive complexity. The equipment may well be constructed with plug-in capability such that the same basic cross-correlator, threshold circuits, and digital logic can be used for a wide variety of tactical screening problems. Thus, the first layer feature-template weights (see Figure 5-1), the second layer decision weights, and the output code used to record the identifications should all be subject to change by means of plug-in modules. In the case of the optically tapped delay line cross-correlator, the first layer weights are readily changed by changing a number of photographic slides. The second layer weights may involve 40 weights per object class per angle of rotation. A separate plug-in resistor matrix for each angle of rotation would represent a practical solution to the problem.

When the size of all target objects for screening are roughly the same, it is necessary only to select a single scanning pattern size determined by the scale of the photography to be processed. When the altitude of flight is approximately fixed for the roll of film to be processed, the scale factor can be set at the console. If, however, altitude is expected to vary significantly over the roll, then it becomes necessary to mark scale factor at frequent intervals on the margin of the film. This can be done manually using a simple digital code punched into the edge of the film, or it may be done automatically during the flight. In either case, the punched marking can be read off the film edge automatically and used to set the scale factor of scan.

If H is the greatest scanning size to be used, then other scanning sizes are chosen to be integer fractions of H ($H/_2$, $H/_3$). For the larger scans, correspondingly larger horizontal pitches of the sawtooth scan and larger spot sizes are also used. During any particular scan, only those object recognition circuits pertaining to that particular scale factor are permitted to record recognition decisions. If scales for scanning vary by factors of two, the increased time required for all scales larger than the smallest is only 4/3 that required for the smallest. When the height and pitch are doubled, the density of picture element samples is 1/4 of the original value, and

$$\sum_{n=0}^{\infty} \left(\frac{1}{4}\right)^n = \frac{4}{3}$$

Depending upon the amount of imagery to be processed and the reaction time required, additional scans may or may not be desirable. Therefore, it should be possible for the equipment to process photography in two scanning modes: a fast mode in which each scale factor is covered only once, and a slow mode in which the image is scanned several times at each major scale factor. The choice between these modes will, of course, be made from the control console as determined by the photo-interpreter.

If, due to economic compromise, the number of scanning lines and video sampling points is less than the theoretical minimum required, a kind of noise results in the recognition circuits (sometimes called quantization noise) which produces different recognition results on nearly identical scans. The effect of this noise is reduced by taking several scans at incrementally different sizes and averaging the outputs of the discriminant function.

The scale for particular photography is known only within a specified range. Scans at several incrementally different scale factors will insure coverage of the scale factor at which the recognition logic most nearly matches

the input imagery. This procedure of size-search also helps to compensate for slight variations in object size.

To reduce film processing time, the equipment could be programmed on the detection of cloud cover, to sample several other areas of the frame rapidly for cloud cover using only the largest scanning frame. If all samples were recognized as cloud cover, the equipment would proceed to the next frame without further scanning. A similar decision could be made for water. Each of these modes could be used or not used according to the decision of the individual operating the equipment. For example, when searching for ships, the water-reject mode would not be appropriate. Selection of mode of operation is made from the control console.

5.7 Estimate of System Complexity

No firm estimate of complexity of an effective image screening equipment can be made until a number of critical experiments have been completed. These experiments are required to determine the following key parameters

- The number of stored picture elements necessary to achieve recognition; present estimate is 2000.
- Number of picture elements which need to be weighted per feature; present estimate is 10 to 50.
- Number of separate feature detectors needed to recognize an object class in one small range of angles; present estimate is 40.
- Number of duplicate feature sets needed to account for rotation; present estimate is 18.
- Number of features required to represent adequately a vocabulary of 20 objects; present estimate is 6 times that required for one object.

In spite of the fact that the foregoing parameters have yet to be determined more closely, a preliminary estimate of complexity based on the above figures may be developed. Consider an image screening system using optically tapped acoustic delay lines and capable of identifying 20 objects of tactical significance. The equipment required to realize such a system is shown in summary form in Table 5-1, "Estimate of Screening System Complexity."

TABLE 5-1

ESTIMATE OF SCREENING SYSTEM COMPLEXITY

Requirement or Assumption	Equipment
200 μsec delay line, 0.1 μsec rise time. Use two:one for gradient, one for gray-scale	2 delay lines
4320 feature detectors [6 x 18 x 40] Each feature detector has 5 transistors, 1 photomultiplier, 50 optical fibers	21,600 transistors 4,320 photomultipliers 216,000 optical fibers
$10 \cdot 9 + 10 \cdot 18 = 270$ Discriminant functions are required on 40 pairs of variables. 10,800 AND gates, 3 diodes each	32,400 diodes
270 thresholds, 5 transistors each	1,350 transistors
Ten 9-legged OR gates plus Ten 18-legged OR gates	270 diodes
Miscellaneous scanning and coding logic	1,000 transistors
Viewing monitor; 32 x 32 light bulb matrix in five colors plus 1,000 transistor logic	5,120 miniature bulbs 320 relays 1,000 transistors

SECTION 6

STATISTICAL METHODS FOR PATTERN CLASSIFICATION

6.1 Introduction

This section presents the main points concerning statistical classification procedures and procedures for examining recorded data from the point of view of significance of predictor variables and configuration of groups in an N-dimensional space. These items are explained in greater detail in Appendix H.

Intuitive classification procedures, based on concepts of distances and directions, and employing transformations of the coordinate space or projections of the samples along a particular direction, were used from the very beginning of multivariate analysis. In 1939, a probabilistic theory of classification first appeared when attention was focused on procedures which would minimize the probability of misclassification. Present classification procedures represent a synthesis of ideas of distance functions or metrics, with probabilistic ideas such as minimizing the probability of misclassification or minimizing the expected loss of misclassification.[1,2,3,4]

The basic assumption underlying probabilistic classification is that there exists for each group π_g, $g = 1, 2, \ldots, G$, a multivariate probability distribution $F_g(x_1, x_2, \ldots, x_N)$. Members of a pattern class are then treated as samples from a population, which are distributed in N-dimensional space according to the distribution associated with that population. This theoretical framework leads to three types of problems which encompass situations ranging from complete statistical knowledge of the distributions, to no knowledge except that which can be inferred from samples. The various situations are discussed individually in the following text.

---

1. C.R. Rao, Advanced Statistical Methods in Biometric Research, John Wiley and Sons, New York, 1952.

2. S.S. Wilks, "Multidimensional Statistical Scatter," Contributions to Probability and Statistics in Honor of Harold Hottelling, Stanford/U. Press, 1960.

3. Maurice M. Tatsuoka, David V. Tiedman, "Discriminant Analysis," Rev. Education Research, Vol 24, pp 402-420, 1954.

4. R.G. Miller, "Statistical Prediction by Discriminant Analysis," Meteorological Monographs, Vol 4, No. 25, October 1962, Amer. Meteorological Society, Boston, Mass.

6.2 Case of Known Distributions

Assuming that the probability density functions $f_g(x_1, x_2, \ldots, x_N)$, $g = 1, 2, \ldots, G$, are known, having a new pattern to classify is the same as having a new observation on the stochastic variable $X = (X_1, X_2, \ldots, X_N)$. One must decide from which of the G distributions this particular observation arose. Several variations can occur within this case.

The simplest situation occurs when one has only two groups whose probibility density functions $f_1(x)$ and $f_2(x)$ are known, and the new pattern must belong to one of these two groups.

In this situation optimal classification is obtained by using the likelihood ratio $L(x) = f_1(x)/f_2(x)$. The value of $x = (x_1, x_2, \ldots, x_N)$ for a new pattern is substituted in $L(x)$ and the result compared with a threshold t. If $L(x)$ exceeds t, the new pattern is classified as belonging to Group 1. Otherwise, the pattern is classified into Group 2. The choice of the threshold t depends on:

1. The criterion of optimality being used.
2. The degree of knowledge of the a priori distribution of the density functions, i.e., the proportion $q_1 : q_2$, $q_1 + q_2 = 1$, of the two groups in the universe from which patterns are drawn.
3. The costs of correct and incorrect classification.

Some of the decision criteria, such as Bayes and Maximum Likelihood, which have been used to obtain values for the threshold t are presented in Appendix H.

If $f_1(x)$ and $f_2(x)$ are multivariate normal distributions with different mean (column) vectors,

$$M^{(g)} = (m_1^{(g)}, m_2^{(g)}, \ldots, m_N^{(g)}), \quad g = 1, 2,$$

and covariance matrices

$$V_g, \quad g = 1, 2,$$

then the logarithm of the likelihood ratio leads to a linear function of the observables when

$$V_1 = V_2$$

and to a quadratic function of the observables when

$$V_1 \neq V_2 .$$

Thus when

$$V_1 = V_2 = V,$$

one has in matrix notation:

$$\log L(x) = x' V^{-1} (M^{(1)} - M^{(2)}) - 1/2 (M^{(1)} + M^{(2)})' V^{-1} (M^{(1)} - M^{(2)}) \quad (6-1)$$

where x' represents the transpose of the column vector $x = (x_1, x_2, \ldots, x_N)$ of measurements made on the pattern which is to be classified;

V^{-1} is the inverse matrix corresponding to the (common) covariance matrix V, which is a N x N matrix with elements

$$v_{ij} = E [(x_i - m_i) (x_j - m_j)] ;$$

$M^{(1)}$ is the mean vector for Group 1;

$M^{(2)}$ is the mean vector for Group 2.

If one lets:

$$C = \log t + 1/2 (M^{(1)} + M^{(2)})' V^{-1} (M^{(1)} - M^{(2)}) \quad (6-2$$

the likelihood ratio principle in this case leads to the classification of the new pattern into one of two regions R_1 and R_2 in the N-dimensional space separated by a hyperplane given by

$$x' V^{-1} (M^{(1)} - M^{(2)}) = C$$

or in expanded form by the equation

$$\sum_{i=1}^{N} (v^{1i} d_1 + v^{2i} d_2 + \ldots v^{Ni} d_N) x_i = C \qquad (6\text{-}3)$$

where v^{ji}, $j = 1, 2, \ldots, N$, are the elements of the matrix V, and $d_j = m_j^{(1)} - m_j^{(2)}$, the difference between the mean values of x_j in the two groups. If one lets:

$$a_i = v^{1i} d_1 + v^{2i} d_2 + \ldots + v^{Ni} d_N \qquad (6\text{-}4)$$

then Equation 6-3 can be rewritten as

$$\sum_{i=1}^{N} a_i x_i = C . \qquad (6\text{-}5)$$

The constant C in Equation 6-5 and the a_i of Equations 6-4 and 6-5 depend only on the known means and covariances in the two groups. If, when substituting the values of x_i measured on a new pattern in Equation 6-5, the left-hand side exceeds C, the pattern in classified into Group 1. If the left-hand side is less than C, the pattern is classified as belonging to Group 2. If the left-hand side equals C, the decision is arbitrary.

If the covariance matrices for the two groups are not equal, i.e., $V_1 \neq V_2$, then instead of the hyperplane of Equation 6-5, the optimum decision boundary, given by the surface of a constant likelihood ratio is:

$$\sum_{i=1}^{N} \sum_{j=1}^{N} \left\{ v_1^{ij} (x_i - m_i^{(1)}) (x_j - m_j^{(1)}) - v_2^{ij} (x_i - m_i^{(2)}) (x_j - m_j^{(2)}) \right\}$$

$$= \text{constant} \qquad (6\text{-}6)$$

where v_1^{ij} and v_2^{ij} are the elements of V_1^{-1} and V_2^{-1} respectively. Equation 6-6 may be written in the form

$$\sum_{\substack{i=1 \\ j=1}}^{N} a_{ij} x_i x_j + \sum_{i=1}^{N} b_i x_i = \text{constant} \tag{6-7}$$

where

$$a_{ij} = \left\{ v_1^{ij} - v_2^{ij} \right\} , \tag{6-8}$$

and

$$b_i = 2 \left\{ (m_1^{(1)} v_1^{il} - m_1^{(2)} v_2^{il}) + (m_2^{(1)} v_1^{i2} - m_2^{(2)} v_2^{i2}) + \right.$$
$$\left. \ldots + (m_N^{(1)} v_1^{iN} - m_N^{(2)} v_2^{iN}) \right\} . \tag{6-9}$$

The expressions of Equations 6-5 and 6-7 will be referred to as linear and quadratic classification functions. They can be shown to be theoretically optimal, with different values for the coefficients, for a number of types of density functions, in addition to normal density functions. Interest in linear and quadratic classification functions stems also from considering them as first- and second-order approximations to arbitrary likelihood ratios, since in many situations they represent the most that can be realized in hardware, or by computation. As an example,[1] for the case where the x_i are binary variables, it is shown in Appendix H that the likelihood ratio of two distributions of N binary variables becomes:

---

1. R. R. Bahadur, "On Classification Based on Responses to n Dichotomous Items," USAF SAM Series in Statistics, Randolph AFB Texas, 1959; also appears in Studies in Item Analysis and Prediction, Herbert Solomon (ed.), Stanford University Press, 1961.

$$L(x) = \frac{\prod_{i=1}^{N} \left(m_i^{(1)}\right)^{x_i} \left(1 - m_i^{(1)}\right)^{(1 - x_i)}}{\prod_{i=1}^{N} \left(m_i^{(2)}\right)^{x_i} \left(1 - m_i^{(2)}\right)^{(1 - x_i)}} \cdot$$

$$\frac{\left[1 + \sum_{i<j} r_{ij}^{(1)} y_i^{(1)} y_j^{(1)} + \sum_{i<j<k} r_{ijk}^{(1)} y_i^{(1)} y_j^{(1)} y_k^{(1)} + \ldots + r_{12\ldots N}^{(1)} y_1^{(1)} y_2^{(1)} y_N^{(1)}\right]}{\left[1 + \sum_{i<j}^{N} r_{ij}^{(2)} y_i^{(2)} y_j^{(2)} + \sum_{i<j<k} r_{ijk}^{(2)} y_i^{(2)} y_j^{(2)} y_k^{(2)} + \ldots + r_{12\ldots N}^{(2)} y_1^{(2)} y_2^{(2)} \cdots y_N^{(2)}\right]} \qquad (6\text{-}10)$$

$$m_i^{(g)} = \text{prob.}\ (x_i^{(g)} = 1)\ ;\ 0 < m_i^{(g)} < 1\ ; \qquad g = 1, 2\ .$$

$$y_i^{(g)} = (x_i - m_i^{(g)}) \Big/ \sqrt{m_i^{(g)} (1 - m_i^{(g)})}\ ; \qquad i = 1, 2 \ldots N;\ g = 1, 2.$$

$$r_{ij}^{(g)} = E_g\, (y_i^{(g)}\, y_j^{(g)})\ , \quad i < j\ .$$

$$r_{ijk}^{(g)} = E_g\, (y_i^{(g)}\, y_j^{(g)}\, y_k^{(g)})\ , \ i < j < k\ . \qquad (6\text{-}11)$$

$$r_{12\ldots N}^{(g)} = E_g\, (y_1^{(g)}\, y_2^{(g)} \ldots y_N^{(g)})$$

with E_g denoting the expectation taken with respect to the probability function of group g, g = 1,2. The r_{ij}, r_{ijk} ... are the "correlation parameters." A second order approximation to the logarithm of L(x) of Equation 6-10 (see Appendix H) is

$$\log L(x) = \sum_{i=1}^{N} (a_i x_i + c_i) + \log \left[1 + \sum_{i<j} r_{ij}^{(1)} y_i^{(1)} y_j^{(1)}\right] - \log \left[1 + \sum_{i<j} r_{ij}^{(2)} y_i^{(2)} y_j^{(2)}\right]\ , \qquad (6\text{-}12)$$

where

$$a_i = \log \left[\frac{m_i^{(1)} \ (1 - m_i^{(2)})}{m_i^{(2)} \ (1 - m_i^{(1)})} \right]$$

and

$$c_i = \log \frac{(1 - m_i^{(1)})}{(1 - m_i^{(2)})} \quad .$$

when the "correlation parameters" are small one can obtain the further approximation

$$\log L(x) = \sum_{i<j} a_{ij} x_i x_j + \sum_{i=1}^{N} b_i x_i \qquad (6\text{-}13)$$

where

$$a_{ij} = u_{ij}^{(1)} - u_{ij}^{(2)}$$

with

$$u_{ij}^{(g)} = r_{ij}^{(g)} \Big/ \sqrt{m_i^{(g)} \ (1 - m_i^{(g)}) \ m_j^{(g)} \ (1 - m_j^{(g)})} \quad , \ g = 1,2;$$

and

$$b_i = a_i + \sum_{j \neq i} (- m_j^{(1)} \ u_{ij}^{(1)} + m_j^{(2)} \ u_{ij}^{(2)}) \ ,$$

with

$$a_i = \log \left[\frac{m_i^{(1)}}{m_i^{(2)}} \cdot \frac{(1 - m_i^{(2)})}{(1 - m_i^{(1)})} \right] \quad .$$

From Equation 6-10 it is apparent that if the number of variables N is moderately large, a second order approximation as represented by Equations 6-12 or 6-13, is as much as one could hope for in general, in the case of binary variables. This does not, of course, prevent consideration of selected higher order correlations in the approximation.

For the case where we have more than two groups, if a priori probabilities q_g and costs of classification k_{jg} are known, then according to the Bayes' criterion, the G functions

$$R(j) = \sum_{g=1}^{G} f_g(x) \cdot q_g \cdot k_{jg}, \quad j = 1,2 \ldots G. \tag{6-14}$$

need to be computed. The new pattern is assigned to that group j for which R(j) is smallest. If all a priori probabilities and costs are taken to be equal, then for a given x, one need only compare $f_g(x)$, g = 1, 2, ..., G and pick that group g for which $f_g(x)$ is largest. In general, partitioning the N-dimensional space into G regions R_g, g = 1, 2, ..., G requires the determination of G(G-1)/2 boundaries given by one likelihood ratio for each pair of groups.

6.3 Case of Parametric Families of Distribution

If a classification procedure is optimal with respect to certain criteria when the functional form and all parameters of the distribution are known, then how "good" are parameters estimated from samples, and what sort of estimates should be used to obtain a good classification procedure?

The performance of a classification procedure using estimated parameters should be consistent with, i.e., tend to, the optimal performance obtained when true parameter values are used. Similarly, estimators are said to be consistent if the estimated value of a parameter approaches the true value of the parameter with probability 1 as the sample size is increased indefinitely. The consistency property, as defined here, will be referenced frequently in the discussion which follows.

For classification into one of two groups, it has been shown[1] that if $f_1(x, \theta)$ and $f_2(x, \theta)$ satisfy certain weak restrictions as to continuity with respect to the parameter θ, then using consistent estimates instead of true parameter values in the likelihood ratio procedure will lead to a consistent classification procedure. For classification of a multivariate observation into one of G groups whose density functions are known except for a number

1. E. Fix, J. L. Hodges, "Discriminatory Analysis: nonparametric discrimination: consistency properties," Report No. 4, USAF School of Aviation Medicine, Randolph Field, Texas, 1951.

of parameters, if consistent estimators are used for the a priori probabilities q_g and for the unknown parameters, it has been shown[1] that the Bayes procedure, with unknown parameters replaced by estimates, is a consistent classification procedure.

When sample sizes are finite, and especially when sample sizes are small, the use of maximum likelihood estimates of parameters in a likelihood ratio procedure or Bayes procedure, in place of the unknown true parameter values can only be justified on heuristic grounds. Nevertheless it is usually done. Thus for the case of two normal density functions with unknown mean vectors and covariance matrices, one has, instead of Equation 6-1, the following expression:

$$\log L(x) = x'S^{-1}(\bar{x}^{(1)} - \bar{x}^{(2)}) - 1/2\,(\bar{x}^{(1)} + \bar{x}^{(2)})\,S^{-1}(\bar{x}^{(1)} - \bar{x}^{(2)}), \qquad (6\text{-}15)$$

where

$$\bar{x}^{(g)} = (\bar{x}_1^{(g)}, \bar{x}_2^{(g)}, \ldots, \bar{x}_N^{(g)}), \; g = 1,2; \qquad (6\text{-}16)$$

$$\bar{x}_i^{(g)} = \frac{1}{n_g} \sum_{r=1}^{n_g} x_{ir}^{(g)}, \; i = 1,2 \ldots, N; \; g = 1,2, \qquad (6\text{-}17)$$

and the elements of S, the estimate of the covariance matrix V, assumed equal for the two groups is obtained by pooling samples from the two groups. The elements of S are given by:

$$s_{ij} = \frac{\sum_{r=1}^{n_1} (x_{ir}^{(1)} - \bar{x}_i^{(1)})(x_{jr}^{(1)} - \bar{x}_j^{(1)}) + \sum_{r=1}^{n_2} (x_{ir}^{(2)} - \bar{x}_i^{(2)})(x_{jr}^{(2)} - \bar{x}_j^{(2)})}{n_1 + n_2 - 2} \qquad (6\text{-}18)$$

For the multigroup case, if the costs k_{jg} are known or can be reasonably assigned, the expression in Equation 6-14 becomes

$$R'(j) = \sum_{g=1}^{G} \hat{f}_g(x) \cdot \hat{q}_g \cdot k_{jg}; \; j = 1,2, \ldots, G \qquad (6\text{-}19)$$

1. P. G. Hoel and R. P. Peterson, "A Solution to the Problem of Optimum Classification," Ann. Math. Statistics, Vol 20, 1949.

where

$$\hat{q}_g = \frac{n_g}{\sum_{g=1}^{G} n_g} ; \; g = 1, 2, \ldots, G. \tag{6-20}$$

and $\hat{f}_g(x)$ is obtained by substituting maximum likelihood sample estimates for true parameters in $f_g(x)$. Thus, for example, if the $f_g(x)$ are assumed to be multivariate normal density functions with equal covariance matrices,

$$\hat{f}_g(x) = \frac{\left|S^{-1}\right|^{1/2}}{(2\pi)^{N/2}} \exp\left[-\frac{1}{2} (x - \bar{x}^{(g)})' S^{-1} (x - \bar{x}^{(g)}) \right] , \tag{6-21}$$

where $\bar{x}^{(g)}$ represents the (column) vector given in Equation 6-16 with elements given by Equation 6-17 for g = 1, 2, . . . , G, and S^{-1} represents the inverse of the matrix S, which by analogy with Equation 6-18 has elements

$$\left(\sum_{g=1}^{G} n_g - G \right) s_{ij} = \sum_{g=1}^{G} \sum_{r=1}^{n_g} (x_{ir}^{(g)} - \bar{x}_i^{(g)}) (x_{jr}^{(g)} - \bar{x}_j^{(g)}) , \tag{6-22}$$

$$i, j = 1, 2, \ldots, N .$$

If the covariance matrices are not assumed to be equal for the various groups, then, instead of obtaining one estimate by pooling all samples it will be necessary to obtain separate estimates S_g with elements

$$s_{ij}^{(g)} = \frac{1}{n_g - 1} \sum_{r=1}^{n_g} (x_{ir}^{(g)} - \bar{x}_i^{(g)}) (x_{jr}^{(g)} - \bar{x}_j^{(g)}) ; \; g = 1, 2, \ldots, G. \tag{6-2}$$

When the forms assumed for the density functions $f_g(x)$ are such that maximum likelihood estimates of parameters are computationally difficult to obtain, the method of moments is used to estimate parameters from sample moments. This method does not posses the optimum property employed by the

maximum likelihood estimator, but may have to be resorted to in some situations, as for example when the functional forms for $f_g(x)$ are assumed to be multivariate versions of the Pearson system of frequency curves.[1, 2]

---

1. M. G. Kendall, The Advanced Theory of Statistics, Vols 1 and 2, Hafner Publishing Co., 1948.

2. W. P. Elderton, Frequency Curves and Correlation, Cambridge University Press, 1938.

6.4 The Nonparametric Case

When the functional forms of $f_g(x)$ are not known, instead of estimating parameters, one must estimate the conditional group probabilities directly. The results obtained with such methods are, of course, inferior to a likelihood ratio procedure using known density functions. The nonparametric case represents the usual practical situation in many problems of classification. It also represents the case which, because of obvious difficulties, has received the least attention to date.

Having available n_g samples from group g, $g = 1, 2, \ldots, G$, one can consider procedures for determining which of a number of assumed forms for the distribution $f_g(x)$ best fits the samples from group g. Directly estimating $f_g(x)$, which for a given new observation x represents real numbers, and comparing a new pattern with known samples from each group, one may determine which set of samples the new pattern must resemble closely.

Here it is appropriate to consider a nonparametric procedure presented by Fix and Hodges [1] for directly estimating the value of the density functions $f_g(x)$ at a new observation $x = (x_1, x_2, \ldots, x_N)$.

Consider the case of two groups. Since substitution of a new observation into the density functions results in the real numbers $f_g(x)$, $g = 1, 2$, once estimates of these two real numbers are obtained they may be substituted into the likelihood ratio. Fix and Hodges present a procedure which provides consistent estimates for the two real numbers $f_g(x)$. The likelihood ratio procedure using these estimates provides consistent discrimination only if the sample sizes are large. For practical situations they suggest the following intuitive procedure: In place of $f_1(x)$ and $f_2(x)$ use Q_1/n_1 and Q_2/n_2 respectively, and compare

$$L = \frac{Q_1/n_1}{Q_2/n_2}$$

with a threshold, t. Here n_g, $g = 1, 2$ represents the sample size in group 1 and group 2, and Q_1 and Q_2 are obtained as follows:

---

1. E. Fix, J. L. Hodges, "Discriminatory Analysis: nonparametric discrimination: consistency properties," Report No. 4, USAF School of Aviation Medicine, Randolph Field, Texas, 1951.

A positive integer K is selected which is large, but small compared to the size of the sample. An appropriate metric is imposed on the sample space. All samples from all the groups are pooled together and of the K values in the pooled sample which are "nearest" to the new point x, Q_1 represents the number from group 1 and $Q_2 = K - Q_1$ represents the number from group 2. Then Q_1/n_1 and Q_2/n_2 are estimates of $f_1(x)$ and $f_2(x)$ respectively. This method of estimating the real numbers $f_g(x)$ corresponding to a new observation carries over to the multigroup case. Note that the concept of nearest samples is stated in terms of an appropriate metric, but that classification is performed by using estimates in a likelihood ratio or Bayes procedure. This matter is discussed in more detail in the following paragraphs.

6.5 Distance, Direction, and Significance

Distance between a new pattern and the mean of a group, between two groups, between samples of two groups, and between a new pattern and samples of a group is often defined in terms of a Euclidean distance function either in original space of the N predictor variables or in a space obtained by a suitable coordinate transformation. Knowing that a pattern or a sample belongs to one of a number of populations, if we can measure the distance of the pattern or sample from each of the several populations then it is reasonable to assign the pattern to that population from which it is least distant.

A distance function which has received attention in discriminant analysis is the Mahalanobis Generalized Distance denoted by D^2. For the case where there are n_1 samples from one group and n_2 samples from a second group, the generalized distance between the two samples is

$$D_N^2 = \sum_{i=1}^{N} \sum_{j=1}^{N} s^{ji} (\bar{x}_i^{(1)} - \bar{x}_i^{(2)}) (\bar{x}_j^{(1)} - \bar{x}_j^{(2)}) \qquad (6\text{-}24)$$

where

- N represents the number of characteristics measured on each pattern;
- s^{ij} are the elements of S^{-1} which is the inverse of the matrix S whose elements are given by Equation 6-18.

As Rao, previously referenced, has stated, D^2 represents ordinary Euclidean distance in a space defined by a set of oblique axes.

If one defines

$$d_i = (x_i^{(1)} - \bar{x}_i^{(2)}) \tag{6-25}$$

then Equation 6-24 becomes

$$D_N^2 = \sum_{i=1}^{N} \sum_{j=1}^{N} s^{ji} d_i d_j \ . \tag{6-26}$$

If one further defines

$$w_{ij} = \sum_{i=1}^{G} \sum_{r_g=1}^{n_g} (x_{ir}^{(g)} - \bar{x}_i^{(g)}) (x_{jr}^{(g)} - \bar{x}_j^{(g)}) \tag{6-27}$$

and lets d denote the column vector $(d_1, d_2, \ldots, d_n)$, then one can rewrite Equation 6-26 as

$$D_N^2 = (n_1 + n_2 - 2)\, d' W^{-1} d \tag{6-28}$$

d' denotes the transpose of d;

W represents the matrix whose elements are w_{ij}.

W is sometimes called the within-samples scatter matrix. This terminology is explained in Appendix H.

In order to use a statistic such as D^2 it is necessary to know its distribution under the various situations which can occur; viz., when the parameters used in the statistic, such as means and covariances, are true population parameters or sample estimates, these two situations being referred to as the "classical" and "studentized" cases respectively, when the characteristics are independent or correlated and when the populations are actually the same or differ in at least one of the N variables. These latter situations are referred to as the central and noncentral cases respectively. The solution of the distribution problem for tests such as D^2 is a difficult one and all available solutions assume that the samples come from multivariate normal populations with equal covariance matrices. The application of such tests to situations in which these assumptions cannot be reasonably made is thus difficult to justify.

To test if the predictors contain information for discrimination, it is necessary to test the hypothesis that the samples actually come from the same group so that D_N^2 would differ from zero only because of the variability within a group. Thus the null hypothesis is that the expectation of the vector d, defined above, is zero. The statistic [1]

$$\frac{(n_1 + n_2 - N - 1)}{(n_1 + n_2 - 2)} \cdot \frac{n_1 n_2}{n_1 n_2} D_N^2 \tag{6-29}$$

is in this situation, i.e., the central case, distributed according to the variance ratio or F distribution with N and $(n_1 + n_2 - N - 1)$ degrees of freedom. This latter terminology arises from the fact that if s_1^2 is the sample estimate of the population variance of a univariate normal distribution based on a sample of size n, and s_2^2 is an independent estimate based on a sample of size n_2, then the distribution of the ratio

$$F = \frac{\frac{n_1 s_1^2}{n_1 - 1}}{\frac{n_2 s_2^2}{n_2 - 1}} \tag{6-30}$$

is the F distribution with $(n_1 - 1)$ and $(n_2 - 1)$ degrees of freedom.

For testing whether adding Q more predictors to the set of N predictors already in use is going to help in increasing the distance between the two groups, Rao [1] suggests the ratio:

$$R = \frac{1 + \frac{n_1 n_2}{(n_1 + n_2)(n_1 + n_2 - 2)} D_{N+Q}^2}{1 + \frac{n_1 n_2}{(n_1 + n_2)(n_1 + n_2 - 2)} D_N^2} \tag{6-31}$$

---

1. C. R. Rao, Advanced Statistical Methods in Biometric Research, John Wiley and Sons, New York, 1952.

where D^2_{N+Q} is the distance based on N + Q characteristics so that in computing w_{ij} from Equation 6-27, one has i, j = 1, 2 . . . , N + Q When D^2_{N+Q} is of the same order as D^2_N, the ratio R will be approximately 1, while a high ratio would indicate that the Q additional characteristics are contributing to the separation between groups. The test statistic in this case is

$$\frac{(n_1 + n_2 - N - Q - 1)}{Q} \cdot (R - 1) , \qquad (6\text{-}32)$$

which is distributed according to the F distribution with Q and $(n_1 + n_2 - N - Q - 1)$ degrees of freedom.

Let Δ^2_N be the "true" Mahalanobis distance between the two population distributions of the first N variables, i.e., Δ^2_N is the distance one would compute if one actually knew the population mean vectors $M^{(1)} = (m_1^{(1)}, \ldots, m_N^{(1)})$ and $M^{(2)} = (m_1^{(2)}, m_2^{(2)}, \ldots, m_N^{(2)})$ and the covariance matric V, assumed equal for the two distributions. The reason that the difference $(D^2_{N+Q} - D^2_N)$ cannot be tested directly is that the distribution of $(D^2_{N+Q} - D^2_N)$ involves Δ^2_N and since Δ^2_N is not known, an exact test of significance cannot be made.

By setting $D^2_o = 0$ and proceeding in a step-by-step manner, the above test may also be used to screen a set of N predictors to find the Q < N best predictors. The first predictor X_1^* is selected by finding out for which one of the N predictors $X_1, X_2, \ldots, X_N$, the ratio, R, of Equation 6-31 is largest, when in its denominator $D^2_N = D^2_o$ and in its numerator $D^2_{N+Q} = D^2_1$. This largest R is substituted in the statistic of Equation 6-32 with N = 0 and Q = 1. If according to this test X_1^* is significant at the prescribed level of significance, then proceed to determine the second best predictor X_2^* of the remaining (N - 1) predictors by finding out for which one

$$R = \frac{1 + \frac{n_1 n_2}{(n_1 + n_2)(n_1 + n_2 - 2)} D_2^2}{1 + \frac{n_1 n_2}{(n_1 + n_2)(n_1 + n_2 - 2)} D_1^2} \qquad (6\text{-}34)$$

is largest. Again the selection of X_2^* is conditional upon its being significant. In this manner selection is continued until X^*_{Q+1} fails to be statistically

significant. Rao's generalization of D^2 to the case of more than two groups, is described in Appendix H. Its use in testing the significance of predictors and in screening variables is described by Rao[1] and Miller.[2] The relation between D^2 and the symmetric divergence, a general measure of the divergence between two populations, and an approximation to the symmetric divergence for the case of binary variables is also described in Appendix H.

Given n_1 samples from group 1 and n_2 samples from group 2, each group can be represented by points in an N-dimensional Euclidean space. Now the question is whether the two sets of samples can be projected onto a line in such a way that in some sense the separation between projected points belonging to different groups is maximized relative to the separation between projected points belonging to the same group.

Let the direction cosines of the line to be determined be proportional to $(a_1, a_2, \ldots, a_N)$, and let

$$z = \sum_{i=1}^{N} a_i x_i \qquad (6\text{-}35)$$

so that $z_1^{(1)}, z_2^{(1)}, \ldots, z_{n_1}^{(1)}$ represent the projected samples of group 1 and $z_1^{(2)}, z_2^{(2)}, \ldots, z_{n_2}^{(2)}$ represent the projected samples of group 2. Further let $\bar{z}^{(1)}$ and $\bar{z}^{(2)}$ represent the mean values of the two samples of z's and let $\bar{z}$ be the mean of the grand sample obtained by pooling the two sets of samples.

Since the first paper on the subject by Fisher,[3] the direction which has usually been considered is that for which the ratio

---

1. C. R. Rao, Advanced Statistical Methods in Biometric Research, John Wiley and Sons, New York, 1952.

2. R. G. Miller, "Statistical Prediction by Discriminant Analysis," Meteorological Monographs, Vol. 4, No. 25, October 1962, Amer. Meteorological Society, Boston, Mass.

3. R. A. Fisher, "The Use of Multiple Measurements in Taxonomic Problems," Annals of Eugenics 7, p. 179-88, Sept. 1936, appears in R. A. Fisher Contributions to Math. Statistics, John Wiley, N. Y., 1952.

$$\frac{\left[\bar{z}^{(1)} - \bar{z}^{(2)}\right]^2}{E(z - \bar{z})^2} \tag{6-36}$$

is maximized. The resulting linear function z is said to have greatest "variance between samples relative to variance within samples." The direction which achieves this is given in terms of the vector $A = (a_1, a_2, \ldots, a_N)$ by

$$A = \text{const. } S^{-1} (x^{(1)} - \bar{x}^{(2)}), \tag{6-37}$$

so that

$$a_i = s^{1i}d_1 + s^{2i}d_2 + \ldots s^{Ni}d_N \tag{6-38}$$

where s^{ji} are elements of S^{-1} which is the inverse of the matrix S whose elements s_{ij} are given by Equation 6-18 and $d_i = (\bar{x}^{(1)} - \bar{x}_i^{(2)})$. The constant multiplier in Equation 6-37 indicates that any vector parallel to A determined from Equation 6-38 will do just as well. If the coefficient values of Equation 6-38 are used in Equation 6-35 to obtain a classification procedure for a new sample with measurements $(x_1 \ldots x_N)$, by comparing $z = \sum_{i=1}^{N} a_i x_i$ with a threshold c, and classifying the sample into group 1 if $z > c$ and group 2 if $z < c$. Then one will be using Fisher's Linear Discriminant Function.[1] It is clear that this linear function is just the first term of the expression in Equation 6-13 for log L(x). Substituting the value of a_i obtained in Equation 6-38 into Equation 6-36 shows that the maximum value of the ratio of Equation 6-36 is just D_N^2 which was defined in Equation 6-26. It is also easily shown that the direction obtained in Equation 6-37 also results when, for the projected samples one maximizes the between-samples scatter while keeping within-samples scatter fixed or minimize the ratio of within-samples scatter to total scatter.[2] The relation of these various maximization and minimization problems to the ratio of

---

1. R. A. Fisher, "The Use of Multiple Measurements in Taxonomic Problems," Annals of Eugenics 7, p. 179-88, Sept. 1936, appears in R. A. Fisher Contributions to Math. Statistics, John Wiley, N. Y., 1952.

2. S. S. Wilks, "Multidimensional Statistical Scatter," Contributions to Probability and Statistics in Honor of Harold Hottelling, Stanford U. Press, 1960.

Equation 6-36 is shown in Appendix H, where it is also shown that classification based on ordinary Euclidean distance between the projection of a new sample and the projected samples from the two groups is equivalent to using a linear discriminant function.

For the case of G groups, $G > 2$, the above method would require $G(G-1)/2$ directions to be computed, one for each pair of groups. However, corresponding to the study of group difference by the projection of two sets of samples onto a line, is the technique of multiple discriminant functions for the overall study of group difference among more than two groups. The method consists of projecting the G sets of N-dimensional sample, $G < N$, onto a Euclidean space of (G-1) dimensions in such a way that the total scatter of the G sets of projected samples in this space of (G-1) dimensions is maximized relative to the within-samples scatter of the samples. The mechanics of obtaining the coordinates of this "Discriminant Space" of (G-1) dimensions is given in Appendix H. Corresponding to each coordinate of the (G-1) dimensional space, one obtains a linear discriminant function. The resulting discriminant functions z_j, $j = 1, 2, \ldots (G-1)$, which are linear functions of the original predictors $x_1, x_2, \ldots, x_N$, may be viewed as a reduced set of variables. Classification in this space of reduced variables can proceed by the methods presented under the heading of probabilistic classification.

6.6 Practical Discriminant Analysis

For the case of two groups, the linear and quadratic classification functions which have been discussed represent feasible solutions which can be justified by one of a number of points of view. The linear function given by the first term of Equation 6-15 is, for normal distributions with covariance matrices assumed equal, near-optimum because of its relation to Equation 6-3. This function can also be justified as being the one which for arbitrary distributions leads to the maximization of the ratio of Equation 6-36. The linear functionor hyperplane represents the simplest boundary which can be used to divide the sample space. Quadratic classification functions represent the next level of complexity. The substitution of sample estimates of parameters in Equations 6-8 and 6-9 and their subsequent use in the quadratic function of Equation 6-7 leads to a near-optimum solution when the distributions are multivariate normal with unequal covariance matrices and the sample sizes are large. For other continuous distributions this quadratic function of Equation 6-7 may be considered as a "good" first approximation. For the case where the covariance matrices are not assumed equal an alternative is to compute the Anderson-Bahadur linear function which finds the direction along which, for arbitrary distributions the ratio of the difference between means of projected samples to the sum of the standard deviations of the two sets of projected samples is maximized (see Appendix H).

For binary variables, a second order approximation to the likelihood ratio is obtained from Equation 6-12 and for the case where the variables appear to be only slightly correlated, a further approximation leads to the quadratic function of Equation 6-13.

When the number of predictors is large an initial screening of the variables may be done by using the generalized distance D^2. This results in a reduction of computation since a number of terms involving unselected variables need not be computed. The reduction in computation is especially valuable when the number of groups involved is large.

For more than two groups, G, a transformation of the sets of samples from the original N-dimensional space to a discriminant space of (G-1) dimensions whose coordinates are linear functions of the original N variables enables one to apply classification procedures, both parametric and non-parametric, in the new space. Since even a linear function of binary variables will be a continuous variable, the joint distributions of a new reduced set of variables will be better approximated by a multivariate normal distribution than the joint distributions of the original variables.

Having obtained a classification procedure it is useful to evaluate its performance on new samples in addition to the known samples which were used to design the classification procedures. Two types of statistics are of interest. The first should test whether the procedure being used does better than a suitably defined "pure chance" method. The second statistic should test whether one classification function performs significantly better than another classification function. It is necessary to have such tests since the performance on a small sample of unknowns may not be indicative of true performance. Some statistics which have been suggested for this purpose are presented by Lubin [1] and Miller. [2]

---

1. Lubin, A., "Linear and Non-Linear Discriminating Functions," Brit. Jour. of Psychology (Stat. Sec.) Vol. 3, pp. 90-104, 1950.

2. R. G. Miller, "Statistical Prediction by Discriminant Analysis," Meteorological Monographs, Vol. 4, No. 25, October 1962, Amer. Meteorological Society, Boston, Mass.

Table 6-1 presents the practical statistical techniques available for generating linear and quadratic discriminant functions suitable for classification of unknown samples belonging to one or the other of two populations. A discusion of the relationship of these techniques to some of the current work on pattern recognition networks has been presented in references 1 and 2.

---

1. Kanal, L., "Evaluation of a Class of Pattern Recognition Networks," Biological Prototypes and Synthetic Systems, Vol. 1, pp. 261-269, New York, Plenum Press, 1962.

2. Kanal, L., et al, "Basic Principles of Some Pattern Recognition Systems," Proc. National Electronics Conference, Vol. 18, pp. 279-295, October 1962.

TABLE 6-1

SUMMARY OF BASIC CLASSIFICATION PROCEDURES

Form of Distribution Known or Assumed	Classification Procedure	Form of Resulting Discriminant Function	Comments	Applicable Equations
1. Multivariate normal distributions with common covariance matrices	Logarithm of likelihood ratio	Sum of terms which are linear in the variables - a hyperplane	Optimal when the assumptions hold. Generally used as a "good" linear classification function	6-1 - 6-5 6-15 and 6-18
2. Multivariate normal distributions with different covariance matrices	Logarithm of likelihood ratio	Sum of linear and quadratic terms: difference of two quadratic forms	Optimal when the assumptions hold. A good quadratic classification function for general use.	6-6, 6-9 used with sample estimates of population parameters
3. Multivariate distributions of binary variables	Logarithm of likelihood ratio	Sum involving the logarithm of terms of the Bahadur expansion for the joint distribution of binary variables	Optimal when all terms of the expansion are retained. Approximations to the distributions, by truncating the expansion and retaining only some terms, may not be probability functions. The procedure does not lead to a simple implementation.	6-10, 6-12
4. Multivariate distributions of binary variables	Bahadur's approximation to the log of the likelihood ratio	Sum of linear quadratic and higher order terms with as many terms being retained as desired	Valid approximation only if the correlation terms retained in the approximation of each joint probability function add up to less than 1. Expansion can be limited to linear or linear plus quadratic terms. Has a simple implementation.	6-13

TABLE 6-1

SUMMARY OF BASIC CLASSIFICATION PROCEDURES (Cont)

Form of Distribution Known or Assumed	Classification Procedure	Form of Resulting Discriminant Function	Comments	Applicable Equations
5. None	Fix-Hodges nonparametric method of estimating the numerator and denominator of the likelihood ratio.	Ratio of two numbers	Distance between a new sample and its K nearest neighboring samples must be computed. Ambiguity in definition of K. Provides consistent discrimination if the sample sizes are large.	page 6-12
6. None	Anderson-Bahadur linear function	A linear function of the variables	Computes a linear function for which the ratio of the difference between the means of the projected samples to the sum of the standard deviations of the projected samples is maximized. This turns out to be the linear function, which, within the class of linear functions, is optimum for case 2 of this table.	H-26 to H-34 (Appendix H)
7. None	Fisher's linear function	A linear function of the variables	Computes a linear function for which the ratio of the variance between projected samples to the variance within the projected samples is maximized. When the distributions are as in case 1 above the resulting expression is the same as in case 1 above except for a constant.	H-55 to H-60 (Appendix H)

SECTION 7

DESIGN OF COMPUTER SIMULATION EXPERIMENTS

7.1 General Description

The overall purpose of the computer simulation experiments is the evaluation of techniques for the recognition of targets of military significance in gray-scale aerial photographs. The input data consist of 100 small photograph segments, each containing a tank image and 150 or more photograph segments containing typical terrain background but no tanks. The experiment will deal with the problem of identifying tanks only, but the approach is sufficiently general to be applicable to a variety of other target types.

The images will be converted to digital form and Laplacian processed in the Philco S-2000 general purpose computer into black and white representations showing only the relatively sharp contrast edges. Using a training sample consisting of 50 tank images and 50 non-tank images, the computer will then design a number of feature masks, one for each of several sub-areas of the image space. These feature masks will be sums of weighted binary picture element values, and in some cases, weighted pairs of picture element values. In the mathematical sense, the masks are linear and quadratic discriminant functions. The weights are of the coefficients of these discriminant functions calculated from the resolution element values obtained in the sample (50 tanks and 50 non-tanks), based on the principles of statistical decision theory. The masks will be designed to produce maximum outputs for the tank images and minimum outputs for non-tank images. There are several alternative procedures for calculating coefficients; the optimum procedure for the present problem has not been determined. Consequently, each alternative procedure must be used to design a set of masks and the performance of the different sets must be compared to determine which is best.

After calculating the coefficients, the computer will calculate and print out the mask outputs for each of the 50 tank and 50 non-tank images in the training sample. There will be a print-out for each of the several sets of masks designed according to the alternative statistical procedures. These print-outs of the feature mask outputs will be studied to determine error and false alarm rates for various assumed threshold settings, in order to obtain an idea of the relative discrimination capabilities of the alternative mask sets.

Assuming a set of thresholds giving equal false alarm and false dismissal error rates, a calculation will be made of the binary values of the threshold feature mask outputs for each member of the training sample and the computer will use this data to calculate the coefficients of a final discriminant

function for each mask set. The computer will then calculate and print out the analog value of this final discriminant function for each image of the training sample. Here again, a print-out will be made for the outputs of each alternative set of feature masks. Inspection of this print-out will show the error rates and false alarm rates obtained for various assumed threshold settings on the final decision function. These figures will provide another basis for comparison of alternative mask designs, but not a conclusive one, inasmuch as the comparison will be based only on the performance of the simulated recognition systems on the training sample, i.e., the same set of images previously used to design the feature masks.

The final design phase of the experiment will provide the decision logic for a set of random masks. The computer will first determine the random mask outputs for the training sample by summing up the products of all the binary resolution element values times the binary values of a random mask at the corresponding points in the image space and repeating this operation for all the random masks. The computer will then calculate an optimum discriminant function and threshold (final decision logic) for the random masks.

In the next phase of the experiment, an evaluation and comparison will be made of the performance of each of the alternative system designs on an unknown sample, i.e., a fresh sample set of tank and non-tank images and used in designing the recognition logics. Fifty new tank images and at least 100 fresh non-tank images will be processed by a computer simulation of the recognition process, and the final discriminant function output before thresholding will be printed out for each member of the sample. This step will be repeated for each simulated recognition logic. Inspection of these print-outs will provide estimates of error rates and false alarm rates for various assumed threshold settings with each simulated system design operating on an unknown sample.

False alarm rates are expected to vary considerably, depending on the degree of resemblance of the non-tank image patterns to tank images. A number of selected non-tank samples from different populations will be put through the computer simulation for one or more simulated recognition systems in order to obtain separate estimates of false alarm rates on each of a number of typical backgrounds (e.g., open country, rocky terrain, built-up areas, and hedgerows). This probably will necessitate a repetition of the computer simulation with additional samples of non-tank images.

After completing and evaluating the results of this first series of experiments simulating systems using random masks and optimum feature masks, it is planned to conduct further experiments on a universal feature mask system. This system will use a number of geometrical and perhaps some texture-feature masks, and test for these features over all sub-areas in the image space. These would be "general purpose" features including straight lines and

parallel straight lines of various lengths at each of a number of angles, corners at various angles and circles. The computer will compile statistics on the numbers of each of these features seen in samples of each of a number of different target classes and non-target classes. The decision logic will then classify an unknown image according to the number of each of the features it contains.

An alternative technique differing somewhat from the one just described will also be investigated in these follow-on experiments. Rather than compiling statistics on the number of features throughout the image space, regardless of the location of these features, the image will be arbitrarily divided into a number of blocks and the existence or non-existence of each feature will be determined for each block. This information will then be used as input to a recognition logic based not only on the occurrence of individual features, but on where these features are located in the image space.

The constraints selected to simplify the problem have been chosen so that they do not prohibit a generalization of the results. The generalizations, hopefully, will permit the establishment of rules and techniques which in turn will lead to the design of more complex recognition logic. The simplifications assumed for the present model are summarized in Table 7-1.

The sample imagery for these experiments is taken from two rolls of aerial photography (negative) taken during Army maneuvers at Fort Drum, New York, in August 1960. The scale of this photography is approximately 1:6000. One hundred images of tanks were cropped from contact prints. These print segments were then assembled into a ten-by-ten mosaic with all the tank images oriented in the same direction. One hundred other photograph segments, containing miscellaneous natural and man-made objects, were similarly assembled into a ten-by-ten mosaic. The mosaics were then photographed and printed as positive transparencies on four-by-five inch glass plates suitable for input to the Philco IMITAC* equipment. The tank mosaics are shown in Figure 7-1. It is planned to prepare additional non-tank mosaics containing selected patterns representative of various typical backgrounds (e.g., open country, built-up areas).

---

* "IMage Input To Automatic Computer" - - a flying spot scanner and associated circuitry that converts an image into digital form to six-bit precision (on tape) preparatory to processing the video data in the Philco S-2000 general purpose computer.

TABLE 7-1

SIMPLIFYING ASSUMPTIONS FOR SIMULATION MODEL

Simplifying Assumption	Comment
Gray-level information is discarded in the pre-processing operation.	Although information is discarded in the process, this technique does not compromise proof-of-feasibility with gray-scale images.
Tank mosaics are presented in registry.	No real performance compromise since registry testing is expected to be used to eliminate registry problems in any future machine or alternately, redundancy is introduced to handle registry.
Interdependence of elements of different regions of the retinal space is ignored in designing masks.	Final decision logic will consider pairwise dependent relationships and thus compensate in part for the simplification.
Initial size of region for which mask is designed is taken as about 50 retinal elements.	Some overlapping of regions will be performed to gain insight into the effect of this limitation.
Restricted population of images.	Careful choice of images to avoid simple cases or unusual views should give a "typical" set.

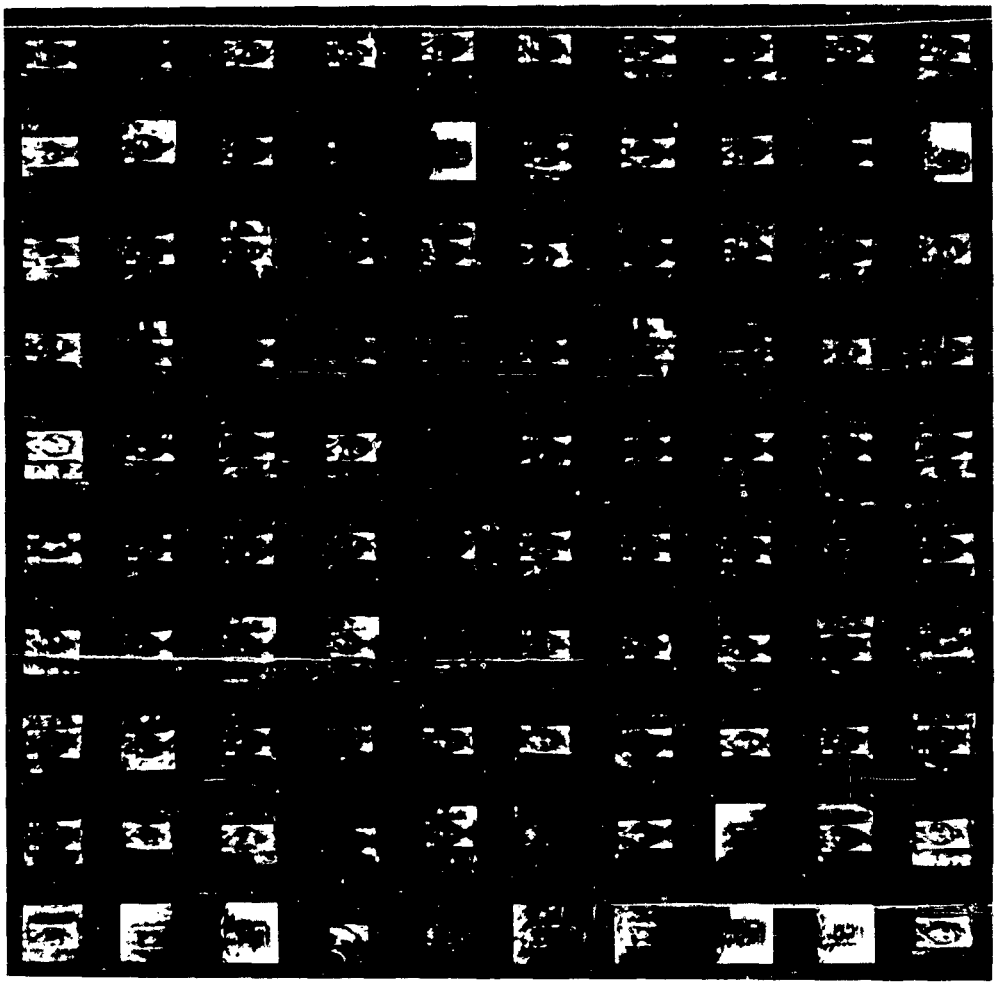


Figure 7-1 Tank Mosaic

7.2 Discriminant Functions Utilized in the Simulation

Introduction

The first requirement is to generate intermediate decision statistics of the form

$$Y_r^{(s)} = \sum_{i=1}^{M_1} b_i x_i + \sum_{i,j=1}^{M_2} a_{ij} x_i x_j \qquad \begin{matrix}(r = 1, \ldots, 6)\\(s = 1, 2, 3)\end{matrix} \quad (7\text{-}1)$$

where,

x_i = the response from the i^{th} retinal element

M_1 = the number of retinal elements in the r^{th} region with significant linear weights, b_i

M_2 = the number of pairs of retinal elements with significant weights, a_{ij}

r = the index indicating the r^{th} subset of retinal elements

s = an index signifying the statistical method applied in determining the weighting coefficients

$Y_r^{(s)}$ = the response from the r^{th} feature mask derived by an application of the s^{th} discriminant function.

These decision statistics (feature masks) will be designed by determining weighting coefficients for each element and each pair of the 50 retinal elements in the r^{th} subset.

Linear Discriminant Function per R. A. Fisher[1]

The linear discriminant function of R. A. Fisher is one of the basic techniques that will be used. Since this is a linear function, Equation (7-1) reduces to

1. R. A. Fisher, "The Use of Multiple Measurements in Taxonomic Problems," Annals of Eugenics 7, p. 179-88, Sept. 1936, appears in R. A. Fisher Contributions to Math. Statistics, John Wiley, N. Y., 1952.

$$Y_r^{(s)} = \sum_{i=1}^{M_1} b_i x_i \quad . \tag{7-2}$$

In this case, M_1 of the linear weights (where $M_1 \leq 50$), will be used to implement the feature mask for the particular region of the retinal space involved. Obviously, the M_1 most significant weights will be selected. It is expected that the Mahalanobis D^2 test of significance described in Section 6 will be satisfactory for selecting the M_1 weights.

Quadratic Discriminant Function with Assumed Normality[1]

Two quadratic discriminant functions are used. They differ in the following general way. The first technique (title above) assumes that the retinal responses, which are binary variables, satisfy a multivariate normal distribution. Thus, the logarithm of the likelihood ratio can be established and weighting coefficients are derived by grouping linear and quadratic terms. Naturally, the covariance matrix of the two populations are assumed to be unequal. Keeping significant coefficients, a quadratic discriminant function of the form of Equation 7-1 is established.

Quadratic Discriminant Function per Lazarsfeld and Bahadur[2]

This second quadratic function makes particular use of the fact that the underlying variables are binary. The actual density function of the retinal responses is approximated by the Lazarsfeld-Bahadur expansion of density functions of binary variables. Dependence higher than pairwise is neglected. This approximation also results in a quadratic form like Equation 7-1, under conditions specified in Appendix H.

7.3 Details of the Computation

Six decision statistics of the form indicated in Equation 7-1 will be generated for each of the three discriminant functions listed above. The computer will calculate sample values of the $Y_r^{(s)}$'s and will print out analog values to be used for evaluation and comparison. This print-out will be utilized to estimate threshold effects, to compare the performance of one set of masks with another, and to judge the effectiveness of the masks for purposes of classification.

1. C.R. Rao, Advanced Statistical Methods in Biometric Research, John Wiley and Sons, New York, 1952.
2. Bahadur, R.R., "On Classification Based on Responses to n Dichotomous Items," USAF SAM series in Statistics, Randolph AFB, Texas, 1959.

Intermediate decisions will be made about the presence or absence of a feature yielding binary variables as inputs to the final decision logic. That is, the final decision statistic for each of the discriminant techniques will take the form

$$Z_s = \sum_{r=1}^{6} c_r Y_r^{(s)}(\eta) + \sum_{r,q=1}^{21} d_{rq} Y_r^{(s)}(\eta) Y_q^{(s)}(\eta) \qquad (7\text{-}3)$$

where

$$Y_r^{(s)}(\eta) = \begin{cases} 1 : Y_r^{(s)} > \eta \\ 0 : Y_r^{(s)} \leq \eta \end{cases} \qquad (7\text{-}4)$$

The coefficients c_r and d_{rq} will be calculated according to the method found most effective in specifying the coefficients a_{ij} and b_i (Equation 7-1 above). The computer will calculate and print out the coefficients c_r and d_{rq} and analog values of Z_s ($s = 1, 2, 3$) for sets of $Y_r^{(s)}$ derived from each discriminant technique. These print-outs will be used in evaluating and comparing these alternative techniques. Final decisions based on whether or not $Z_s > \theta$ will be made on test samples containing 50 line drawings* of tank mosaics and two or more groups of 50 line drawings of non-tank populations.

The results of the above experiment will be used as a guide for the design of feature masks over the entire retinal space. It is expected that some overlapping sub-regions of the retinal space will be permitted in the design of feature masks at this stage of the experiment to provide some insight into the effects of the limited size of the sub-regions. Some 16 to 32 feature masks will be designed in conjunction with a "best" final decision logic. These will be used in the next experiment for comparison to the random mask approach.

The majority of the calculations required of the computer will involve solving the equations listed in the paragraphs below.

Detailed Calculations for Linear Discriminant Function

Solve for b_i, where

$$b_i x_i = \left(\sum_{j=1}^{50} s^{ji} (\overline{x}_j(1) - \overline{x}_j(2)) \right) x_i \qquad (7\text{-}5)$$

* i.e., detail detected representations using the thresholded Laplacian.

or

$$b_i = (S^{1i} d_1 + S^{2i} d_2 + \ldots S^{50i} d_{50}) \tag{7-6}$$

where

$$d_i = (\overline{x}_i^{(1)} - \overline{x}_i^{(2)}) \tag{7-7}$$

and s^{ji} = an element of the inverse transform of S defined in Equation 6-18, Section 6.

Repeat for six different subsets of $\{x\}$ of 50 elements each, for all b_i such that $|b_i| > \frac{k}{50}$ $(k \geq 1)$, compute the statistics

$$Y_{jg} = \sum_{i=1}^{M} b_i x_i \quad \begin{matrix}(j = 1, \ldots, S_g) \\ (g = 1, 2)\end{matrix} \tag{7-8}$$

$$\overline{Y}_1 = \frac{1}{S_1} \sum_{j=1}^{S_1} Y_{j1} \tag{7-9}$$

$$\overline{Y}_2 = \frac{1}{S_2} \sum_{j=1}^{S_2} Y_{j2} \tag{7-10}$$

where

S_1 = number of target samples of class 1

S_2 = number of target samples of class 2

The estimated computer time on the Philco 2000 required to complete computations of all Y_j's, Y_1's, and Y_2's is less than 30 seconds without print-out.

Print-out of all 300 b_i's and all Y_j's, Y_1's, and Y_2's, will be required to provide analog data for engineering analysis and evaluation. An additional 15 seconds of computer time is approximated for print-out.

Calculation for Likelihood Ratio with Assumed Normality

Compute

$$a_{ij} = \left\{ v_1^{ij} - v_2^{ij} \right\} \quad (i, j = 1, \ldots, 50) \tag{7-11}$$

and

$$b_i = 2 \left\{ (m_1^{(1)} v_1^{il} - m_1^{(2)} v_2^{il}) + (m_2^{(1)} v_1^{i2} - m_2^{(2)} v_2^{i2}) + \ldots \right.$$
$$\left. + (m_N^{(1)} v_1^{iN} - m_N^{(2)} v_2^{iN}) \right\} \tag{7-12}$$

Keeping significant values of b_i and a_{ij}, compute

$$Y_{ks} = \sum_{i=1}^{M_1} b_i x_i + \sum_{i,j=1}^{M_2} a_{ij} x_i x_j \qquad \begin{matrix} (s = 1, \ldots, 100) \\ (k = 1, \ldots, 6) \end{matrix} \tag{7-13}$$

where

M_1 = number of significant values of b_i's

M_2 = number of significant pairs, i.e., of a_{ij}'s

s = index over target sample

k = index over features for which masks are required.

With print-out of the sample means, sample covariances, and the coefficients, the automatic design of six masks will required less than one minute of computer time.

Calculation for Second Order Binary Discriminant Function
(Lazarsfeld-Bahadur Expansion)

Compute

$$a_{ij}^{(1)} = u_{ij}^{(1)} - u_{ij}^{(2)} \tag{7-14}$$

where

$$u_{ij}^{(g)} = \frac{r_{ij}^{(g)}}{\sqrt{m_i^{(g)} (1 - m_i^{(g)}) \; m_j^{(g)} (1 - m_j^{(g)})}}, \quad (g = 1, 2) \quad (7\text{-}15)$$

and

$$b_i = a_i + \sum_{j \neq i}^{N} (-m_j^{(1)} u_{ij}^{(1)} + m_j^{(2)} u_{ij}^{(2)}) \quad (7\text{-}16)$$

with

$$a_i = \log \left[\frac{m_i^{(1)}}{m_i^{(2)}} \; \frac{(1 - m_i^{(2)})}{(1 - m_i^{(1)})} \right] \quad (7\text{-}17)$$

(Note: see Section 6 for approximations made in determining these coefficients.)

Keeping significant weighting coefficients, compute six feature mask responses:

$$Y_{ks} = \sum_{i=1}^{M_1} b_i x_i + \sum_{i,j=1}^{M_2} a_{ij} x_i x_j \quad (7\text{-}18)$$

$$(s = 1, \ldots, 100)$$

$$(k = 1, \ldots, 6)$$

as in the likelihood ratio technique described above.

With print-out of the sample means, sample covariances, and the coefficients, the automatic design of six masks will require less than one minute of computer time.

Final Decision Logic and Comparison Experiment

A choice of the discriminant function to be used in the final decision will be made on the basis of experimental results. Each set of six masks

derived by the three methods will be used to generate inputs to the final decision. Scores or frequencies of the two ways of misclassification will be accumulated on the basis of six masks as a function of the threshold of the final decision. With print-out of the final decision statistics for each target sample, computer time for this phase is expected to be from two to five minutes.

These results will be used as a guide for final design of the simulated tank recognition device. From 16 to 32 masks will be designed by the preferred discriminant technique, i.e., an additional 10 to 26 masks. Depending upon the results of previous phases, the required computer time may be as much as one hour or as little as a few minutes.

SECTION 8

BIBLIOGRAPHY

This section is bound separately as Volume II. It is classified SECRET.

SECTION 9

PLANS FOR THE NEXT INTERVAL

Implementation Studies

The problems inherent in parallel access optical correlators will be examined in greater detail. In particular, we will study the feasibility of layering subdecisions using cascaded multimask arrays or mosaics of optical correlators to achieve data rates competitive with multiple cross-correlation techniques based on scanning.

Computer Simulation Experiments

Work will continue on the experiments described in Section 7. Computer programs will be prepared for feature mask design, for generating random mask outputs, and for computing mask and final decision logic outputs. These computer programs will operate on the training sample (fifty tanks and fifty non-tanks) to generate print-outs from which alternative feature masks designs will be evaluated.

Computer simulation of performance of feature masks and random masks on unknown samples will be completed if time permits.

SECTION 10

IDENTIFICATION OF KEY TECHNICAL PERSONNEL

Name	Title	Man - Hours 9/1/62 - 11/30/62	Man - Hours 12/1/62 - 2/28/63
*T. J. B. Shanley	Manager, Recognition Laboratory	indirect labor	
T. J. Harley, Jr.	Research Group Supervisor (Project Manager)	257	411
J. S. Bryan	Research Section Manager (Chief Project Scientist)	148	238
J. B. Chatten	Manager	45	19
*L. Kanal	Manager	93	115
*W. F. Werner	Research Specialist	12	0
*D. R. Taylor	Research Specialist	0	212
*J. Z. Grayum	Research Specialist	0	149
*C. Gumacos	Research Specialist	0	213
*H. G. Kellett	Senior Engineer	0	320
*J. R. Richards	Senior Engineer	48	458
H. H. Schaffer	Senior Engineer	426	40
*J. Mantell	Physicist	0	265
*H. Domabyl	Engineer	0	248

* Biographies of these personnel are included in Appendix I.

APPENDIX A

LAPLACIAN AND GRADIENT PREPROCESSING

1. Introduction

Two candidate preprocessing techniques are presently under detailed consideration for modifying input video data prior to its introduction into an imagery screening cross-correlator. They are: Laplacian (spatial band-pass filtering), and gradient magnitude (spatial differentiation). The Laplacian operation eliminates absolute brightness scale as well as low-spatial frequencies which are of little consequence in screening operations. Gradient magnitude detection is a non-linear operation that responds to edges in the scanned photograph. Each technique has distinct advantages, and either may be implemented simply with short electrical delay lines.

The chief advantage of the Laplacian is that it responds linearly to a large fraction of relevant input data, while the gradient rejects more data and presents the retained data in a form more readily usable in a later cross-correlation process.

Delay line implementations of the two preprocessing techniques are similar in concept. A ribbon scan is assumed (see Section 4.3); Figure A-1 shows part of the path followed by the scanning spot image prior to its arrival at point A on the transparency. Through insertion of proper delays, simultaneous sensing at any number of points A through F is possible. The Laplacian approximation L may be taken as

$$L = V_D - 1/4\ (V_A + V_C + V_E + V_F), \tag{A-1}$$

and the gradient magnitude $|G|$ as

$$|G| = \sqrt{(V_A - V_E)^2 + (V_B - V_D)^2}\ . \tag{A-2}$$

Alternately, the gradient squared may be obtained and

$$|G|^2 = (V_A - V_E)^2 + (V_B - V_D)^2\ . \tag{A-3}$$

The quantities V_A, V_B . . . , are voltages sensed at the points, A, B, . . . , in Figure A-1.

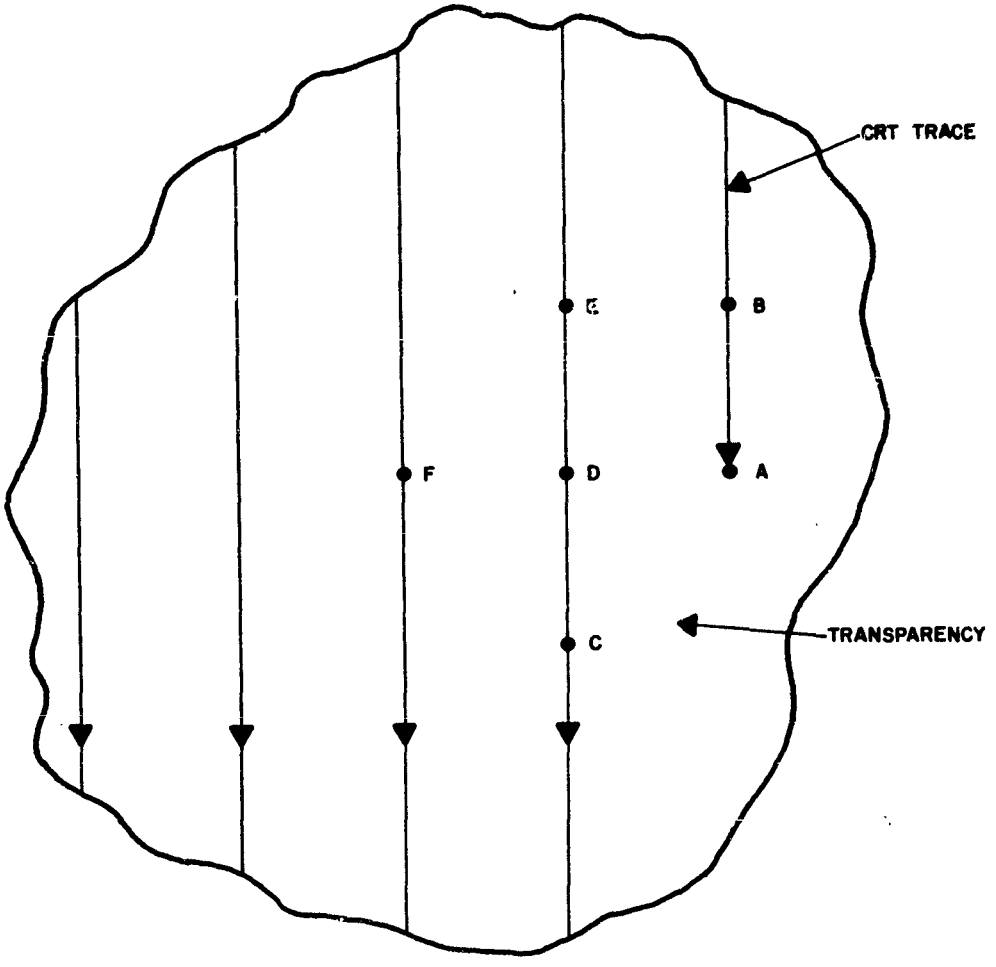


Figure A-1 Gradient and Laplacian Detection

In the following discussion the two delay line preprocessing techniques are examined to determine their signal-to-noise ratio capabilities.

2. Laplacian

Figure A-1 illustrates simultaneous delay line sensing of several data points in an array of fixed form. The Laplacian (L) at sample point D may be obtained approximately by taking

$$L_D = V_D - 1/4\ (V_A + V_C + V_E + V_F). \tag{A-4}$$

The voltage (V) sensed at each point is simply input video delayed by some fixed interval, plus noise which is assumed to be gaussian and independent of that at the other terminals. The output noise component (N_D) due to the Laplacian operation above is simply the weighted rms summation of noise components (rms value σ_i), i.e.,

$$N_D = \sqrt{\sigma_i^2 + \frac{1}{16}\ (4\ \sigma_i^2)} = 1.12\ \sigma_i \tag{A-5}$$

Because the Laplacian signal output differs in form from that at the input, it is necessary to define criteria by which signal-to-noise ratio comparisons may be made. For example, we may take as the input signal, video due to scanning vertically across a horizontal edge and let $(S/N)_{in}$ represent the ratio of the resulting video step to rms noise in the video signal. Since the Laplacian and gradient squared functions will generally be used in conjunction with threshold detectors, it is reasonable to define $(S/N)_{out}$ as the ratio of the peak deviation of the signal from zero to the rms noise in the output. When a horizontal transition lies between points C and D, the instantaneous output signal amplitude is one quarter of the input step. When the transition lies between D and E, the output is one quarter of the input step, but in the opposite polarity. The relationship between output and input signal-to-noise ratio may now be written for this case as

$$(S/N)_{out} = \frac{0.25\ S}{1.12\ \sigma_i} = 0.223\ (S/N)_{in}. \tag{A-6}$$

Similarly, it may be shown for a line of width approximately equal to element spacing that

$$(S/N)_{out} = 0.446\ (S/N)_{in} \tag{A-7}$$

and for a point of about this diameter

$$(S/N)_{out} = 0.892\ (S/N)_{in} \ . \qquad \text{(A-8)}$$

While these figures are obtained for specific inputs, it is believed that they are indicative of performance to be expected from the delay line Laplacian detector.

3. Gradient Squared

The second preprocessing technique to be considered is the gradient (magnitude) squared. By this process, level changes are detected without regard to the direction of the change. The delay line gradient squared detector operates on the points, A, B, D and E in Figure A-1, and the equation of this operation is, (ignoring noise)

$$|G|^2 = (V_A - V_E)^2 + (V_B - V_D)^2 \ . \qquad \text{(A-9)}$$

Since there is a signal component and a noise component associated with each of the four points, it is convenient to rewrite Equation A-9 with noise terms included. The output voltage (V_o) is thus found to be

$$V_o = \left[(V_A - V_E) + (n_A - n_E)\right]^2 + \left[(V_B - V_D) + (n_B - n_D)\right]^2 \qquad \text{(A-10)}$$

$$V_o = \left[(V_A - V_E)^2 + (V_B - V_D)^2\right] + 2\left[(V_A - V_E)(n_A - n_E) + (V_B - V_D)(n_B - n_D)\right] + \left[(n_A - n_E)^2 + (n_B - n_D)^2\right] \qquad \text{(A-11)}$$

$$= V_1\,(S) + V_2\,(S, N) + V_3\,(N) \qquad \text{(A-12)}$$

with $V_1\,(S) = |G|^2$.

$V_2(S,N)$ is a linear signal-weighted noise term whose rms value σ_{V2} is

$$\sigma_{V2} = 2\left[(V_A - V_E)^2(2\sigma_i^2) + (V_B - V_D)^2(2\sigma_i^2)\right]^{1/2} \quad \text{(A-13)}$$

$$= 2\sqrt{2}\sigma_i \left[(V_A - V_E)^2 + (V_B - V_D)^2\right]^{1/2} \quad \text{(A-14)}$$

$$= 2\sqrt{2}\sigma_i \, |G| \quad \text{(A-15)}$$

where σ_i is the standard deviation of noise at each delay line terminal.

$V_3(N)$ is a square law summation of two random noise terms resulting in a chi-square distribution.[1] Since each noise term n_A, n_E n_B, n_D has rms value σ_i, we may let $X_1 = n_A - n_E$ and $X_2 = n_B - n_D$ with $\sigma_{x1} = \sigma_{x2} = \sigma_x\sqrt{2}\sigma_i$. Then the chi-square distribution gives

$$f(V_3) = \frac{1}{2\sigma_x^2} \, e^{-\frac{V_3}{2\sigma_x^2}} \quad \text{(A-16)}$$

The mean (μ_{V3}) and standard deviation (σ_{V3}) of this density function are equal and

$$\mu_{V3} = \sigma_{V3} = 2\sigma_x^2 \quad \text{(A-17)}$$

Since $\sigma_x = \sqrt{2}\sigma_i$,

$$\mu_{V3} = \sigma_{V3} = 4\sigma_i^2 \, . \quad \text{(A-18)}$$

As with the Laplacian, it is necessary to set up conditions by which output signal-to-noise ratio may be compared with that into the gradient detector. We once again consider a horizontal edge with $(S/N)_{in}$ the ratio of this step to rms noise at any tap. It is clear that the output due to this edge is maximum when

1. Modern Probability Theory and its Applications, Emanuel Parzen. John Wiley and Sons, N. Y. 1960, p. 181.

the edge passes through center of the array, and is twice the input step magnitude. Therefore,

$$|G| = \sqrt{2} S_{in} \tag{A-19}$$

$$|G|^2 = 2 S_{in}^2 \tag{A-20}$$

and we now have the ingredients necessary to determine how gradient detection affects signal-to-noise ratio.

It is convenient to obtain expressions for very high and very low signal-to-noise ratios.

a. High Signal-to-Noise Ratio

In this case, the linear term $V_2(S, N)$ of Equation A-12 predominates, and we may write

$$(S/N)_{out} = \frac{2 S_{in}^2}{2\sqrt{2}\sigma_i \sqrt{2} S_{in}} = 0.5 (S/N)_{in} \tag{A-21}$$

b. Low Signal-to-Noise Ratio

The quadratic term $V_3(N)$ of equation (A-9) predominates in this case, and

$$(S/N)_{out} = \frac{2 S_{in}^2}{4 \sigma_i^2} = 0.5 (S/N)_{in}^2 \tag{A-22}$$

In general, any practical system will have a high signal-to-noise ratio, and Equation (A-21) will apply. Thus it can be expected that few signal-to-noise ratio problems should arise in gradient detection.

Since different system concepts may be necessary to utilize the two outputs most effectively, it would be presumptious to attempt an absolute comparison between Laplacian and gradient magnitude in terms of imagery screening capabilit However, so little loss in signal-to-noise ratio results from gradient magnitude detection, and because simplified screening logic may be used with this type of preprocessing, the gradient process appears more attractive at present.

APPENDIX B

EQUIVALENCE OF SPATIAL FREQUENCY FILTERING AND CROSS CORRELATION TECHNIQUES

The coherent light, optical spatial filtering technique provides an exact two-dimensional analog of conventional passive filtering in the one-dimensional time domain. Due to the apparent generality of the technique, it is sometimes assumed to be superior in its capabilities to the techniques that work in the spatial domain, such as the lensless correlographs and scanner, tapped delay-line cross correlators. In fact, these spatial domain techniques are completely equivalent to optical spatial frequency filtering. The mathematical statement relating the spatially filtered output $\bar{f}(x, y)$ to the raw input $r(x, y)$ in the ideal coherent-light, optical spatial filter is as follows.

$$\bar{f}(x, y) = \mathcal{F}^{-1} \left[\bar{F}(\omega_x, \omega_y)\right] \tag{B-1}$$

$$\bar{F}(\omega_x, \omega_y) = \bar{R}(\omega_x, \omega_y) \cdot \bar{G}(\omega_x, \omega_y) \tag{B-2}$$

$$\bar{H}(\omega_x, \omega_y) = \mathcal{F}[(h(x, y)] \tag{B-3}$$

where $\bar{G}(\omega_x, \omega_y)$ is the filter function. In general, $r(x, y)$ will be real, but the other functions may be complex, i.e.,

$$\bar{f}(x, y) = f_r(x, y) + j\, f_i(x, y) \tag{B-4}$$

$$\bar{F}(\omega_x, \omega_y) = F_r(\omega_x, \omega_y) + j\, F_i(\omega_x, \omega_y) \tag{B-5}$$

$$\bar{G}(\omega_x, \omega_y) = G_r(\omega_x, \omega_y) + j\, G_i(\omega_x, \omega_y) \tag{B-6}$$

$$\bar{R}(\omega_x, \omega_y) = R_r(\omega_x, \omega_y) + j\, R_i(\omega_x, \omega_y) \tag{B-7}$$

and $$f(x, y) = f_r(x, y). \tag{B-8}$$

Note that in general, $\bar{G}(\omega_x, \omega_y)$ has an inverse transform

$$\bar{g}(x, y) = \mathcal{F}^{-1}\left[G(\omega_x, \omega_y)\right] = g_r(x, y) + j\, g_i(x, y). \tag{B-9}$$

The output $\bar{f}(x, y)$ can be expressed equally well by the expression

$$\bar{f}(x, y) = r(x, y) * \bar{g}(x, y) \tag{B-10}$$

where * is the symbol for convolution, i.e.,

$$\mathcal{F}\left[r(x, y) * \bar{g}(x, y)\right] = \bar{R}(\omega_x, \omega_y) \cdot \bar{G}(\omega_x, \omega_y) \tag{B-11}$$

The complete expression of the convolution is

$$\bar{f}(x, y) = r(x, y) * \bar{g}(x, y)$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} r(x - \xi, y - \eta)\, \bar{g}(\xi, \eta)\, d\xi\, d\eta\,. \tag{B-12}$$

Equation B-12 is the expression for the cross-correlation of $r(x, y)$ and $\bar{g}(x, y)$, which is the operation performed by the lensless correlograph and by the scanner-delay line cross-correlator techniques.

Throughout the discussion, the variable $\bar{g}(x, y)$ has been assumed to be complex. However, it should be noted that this does not limit the applicability of the cross-correlator techniques which work only with real functions. The complex function $\bar{g}(x, y)$ can be represented by the sum of a real function and an imaginary function.

$$\bar{g}(x, y) = g_r(x, y) + j\, g_i(x, y).^{1} \tag{B-13}$$

The correlation of $r(x, y)$ with $g_r(x, y)$, and of $r(x, y)$ with $g_i(x, y)$, can be carried out separately (but simultaneously in the delay line cross-correlator device). Thus, the same information is available at the cross-correlator output as is available in the complex output of the coherent light optical

---

1. $g_r \neq \mathcal{F}^{-1}\left[G_r(\omega_x, \omega_y)\right]$ and $g_i \neq \mathcal{F}^{-1}\left[G_i(\omega_x, \omega_y)\right]$

spatial filter. Because of the nature of practical light sensors, the output from the optical spatial filter device will be measured in terms of quadratic content

$$\left| \bar{f}(x, y) \right|^2 = \bar{f}(x, y)\, \bar{f}^{\,*}(x, y) = f_r^{\,2}(x, y) + f_i^{\,2}(x, y). \qquad \text{(B-14)}$$

With the cross-correlation techniques, the real and imaginary components can be treated separately.

The foregoing discussion has been carried out in terms of complex aperture functions in order to show completely general equivalence between coherent-optical spatial filtering and cross-correlation. As a practical matter, however, useful apertures for spatial filtering are almost always pure real (no imaginary component). For mechanical convenience, investigators in coherent optical systems often use unsymmetrical filter functions in ω which result in complex apertures; however, there is no evidence that the use of complex apertures in the spatial domain provides any additional utility over pure real apertures.

APPENDIX C

EQUIVALENCE OF TWO TEXTURE MEASUREMENT TECHNIQUES

In this appendix, the equivalence of two methods of measuring "textural" features is demonstrated. In the first method, the output ϕ for a particular texture measurement is obtained by applying a linear discriminant function to the components of the Fourier power spectrum of the image. In the second method, the input image is cross-correlated with each of two apertures, these correlation functions squared, and the difference of these squares integrated over the entire frame, yielding the output ϕ. It is shown that with the appropriate choice of aperture functions, the two expressions for ϕ are equivalent.

Let $b(x, y)$ be the input image, and let $B(\omega_x, \omega_y)$ be its Fourier transform, and $B(\omega_x, \omega_y) \cdot B^*(\omega_x, \omega_y)$ the Fourier power spectrum. $D(\omega_x, \omega_y)$ is a two dimensional linear discriminant function which will operate on $\left| B(\omega_x, \omega_y) \right|^2$.

$$\phi = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} D(\omega_x, \omega_y)\, B(\omega_x, \omega_y)\; B^*(\omega_x, \omega_y)\, d\omega_x\, d\omega_y \qquad \text{(C-1)}$$

Define two real non-negative functions

$$D_p(\omega_x, \omega_y) = \begin{cases} D(\omega_x, \omega_y) \text{ when } D(\omega_x, \omega_y) > 0 \\ 0 \qquad\qquad\qquad \text{otherwise} \end{cases} \qquad \text{(C-2)}$$

and

$$D_N(\omega_x, \omega_y) = \begin{cases} -D(\omega_x, \omega_y) \text{ when } D(\omega_x, \omega_y) \leq 0 \\ 0 \qquad\qquad\qquad \text{otherwise} \end{cases} \qquad \text{(C-3)}$$

where

$$D_p(\omega_x, \omega_y) - D_N(\omega_x, \omega_y) = D(\omega_x, \omega_y)\,. \qquad \text{(C-4)}$$

In addition, define two aperture functions

$$W_1(x, y) = \mathcal{F}^{-1}\left[\sqrt{D_p(\omega_x, \omega_y)}\right] \tag{C-5}$$

$$W_2(x, y) = \mathcal{F}^{-1}\left[\sqrt{D_N(\omega_x, \omega_y)}\,.\right] \tag{C-6}$$

Then let

$$f_1(x, y) = W_1(x, y) * b(x, y) = \mathcal{F}^{-1}\left[\sqrt{D_p(\omega_x, \omega_y)}\,\cdot\, B(\omega_x, \omega_y)\right] \tag{C-7}$$

and

$$f_2(x, y) = W_2(x, y) * b(x, y) = \mathcal{F}^{-1}\left[\sqrt{D_N(\omega_x, \omega_y)}\,\cdot\, B(\omega_x, \omega_y)\right]. \tag{C-8}$$

The total quadratic content (proportional to power) of any function can be obtained by integrating over the function or over its transform.

$$\therefore\ 4\pi^2 \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_1^2(x, y)\, dx\, dy = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} D_p(\omega_x, \omega_y)\, B(\omega_x, \omega_y)\, B^*(\omega_x, \omega_y)\, d\omega_x\, d\omega_y \tag{C-9}$$

and

$$4\pi^2 \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_2^2(x, y)\, dx,\, dy = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} D_N(\omega_x, \omega_y)\, B(\omega_x, \omega_y)\, B^*(\omega_x, \omega_y)\, d\omega_x\, d\omega_y \tag{C-10}$$

The expression for ϕ can be obtained from the combination of Equations (C-4), (C-9), and (C-10) in Equation (C-1).

$$\phi = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left[D_p(\omega_x, \omega_y) - D_N(\omega_x, \omega_y)\right] B(\omega_x, \omega_y)\, B^*(\omega_x, \omega_y)\, d\omega_x\, d\omega_y \tag{C-11}$$

$$\therefore\ \phi = 4\pi^2 \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left[f_1^2(x, y) - f_2^2(x, y)\right] dx\, dy \tag{C-12}$$

Note that for perfect equivalence, the integrals have to be over the range $-\infty$ to ∞. However, in fact, the image data of interest will be limited to a particular frame area, with the remainder of the range blacked out and the aperture functions will be defined over a much smaller area,

$$\therefore \quad \phi = 4\pi^2 \int_{\text{frame}} \int \left[f_1^{\,2}(x, y) - f_2^{\,2}(x, y) \right] dx\, dy \qquad \text{(C-13)}$$

APPENDIX D

OPTICAL EFFICIENCY OF THE LENSLESS CORRELOGRAPH

In order to evaluate the performance of sensors for use in the image plane of the lensless correlograph, it is necessary to determine the light flux density at the image plane (illuminance, E) that will result when the peak value (B) of the luminous emittance (L) is specified at the object plane.

For this derivation, assume that the input photograph is a Lambert surface. The luminous emittance of any point of the photograph is

$$L(x, y) = B\mu_o(x, y),$$

where B is a constant corresponding to the emittance at peak white, and $\mu_o(x, y)$, $0 \leq \mu_o \leq 1$, is a function of location in the image. Consider the light from a single square-picture element of area a_o^2 passing through an aperture with transmittance μ_a, $0 \leq \mu_a \leq 1$, and impinging on a point in the image space (see Figure D-1). The illuminance at that point from this picture element area is given by

$$\Delta E = \frac{\mu_a L a_o^2}{2\pi d^2} = \frac{\mu_o \mu_a B a_o^2}{2\pi d^2} . \qquad \text{(D-1)}$$

The total illuminance at a point in the object plane due to the cross-correlation of a input photograph with a weighted aperture is given by the following summation over all the m elements of the aperture:

$$E = \frac{B a_o^2}{2\pi d^2} \sum_{1}^{m} \mu_o \mu_a . \qquad \text{(D-2)}$$

Let

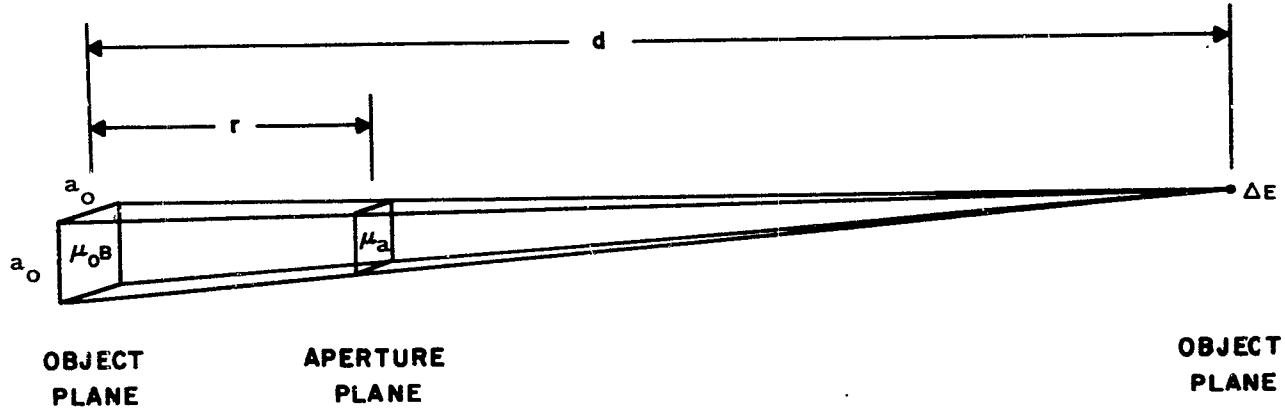
$$\mu = \frac{1}{m} \sum_{1}^{m} \mu_o \mu_a , \qquad \text{(D-3)}$$

Figure shows lightrays converging on a point in the image plane from a single picture element in the object plane. The illuminance, or light flux density, is given by

$$\Delta E = \frac{\mu_o \mu_a B a_o^2}{2\pi d^2} \qquad \text{where}$$

$\mu_o B$ = luminance of input picture element and

μ_a = transmittance of aperture.



$$\text{MAGNIFICATION RATIO} = \frac{d-r}{r} = M$$

$$\text{OR} \quad \frac{r}{d} = \frac{1}{M+1}$$

Figure D-1 Geometry Relating Single Elements of Lenless Correlograph

then

$$E = \frac{m\mu \ B \ a_o^2}{2 \pi \ d^2} \quad . \tag{D-4}$$

If the input photograph is a square, $\sqrt{A_o}$ units along a side, then

$$n = \frac{A_o}{a_o^2} \tag{D-5}$$

is the total number of picture elements, and

$$E = \frac{m\mu \ B \ A_o}{n \ 2 \pi \ d^2} \quad . \tag{D-6}$$

There is a 'rule of thumb' for optical systems designed to eliminate general aberrations that reduce image quality and resolution, which states that the half angle described by the object with respect to a point on the optical axis in the aperture plane should be less than ten degrees.

Therefore, let

$$\sqrt{\frac{A_o}{2}} / r \approx \tan 10^o, \tag{D-7}$$

which yields

$$r^2 \approx 8 \ A_o \tag{D-8}$$

Then

$$E = \frac{m\mu \ B \ r^2}{n \ 16 \pi \ d^2} \tag{D-9}$$

If an output sensor element has a linear dimension, a_i, and the geometry of the correlograph is fixed to match the resolution of the sensor to the

input, then $M = \frac{a_i}{a_o}$, and

$$\frac{r}{d} = \frac{1}{1 + M} = \frac{a_o}{a_i + a_o} , \tag{D-10}$$

then

$$E = \frac{m\mu\ B}{n\ 16\ \pi} \left(\frac{a_o}{a_o + a_i} \right)^2 . \tag{D-11}$$

Consider the distribution of μ_o and μ_a to be uniform between 0 and 1. Then the expected values of these terms and their squares are

$$\mathcal{E}(\mu_o) = \mathcal{E}(\mu_a) = 1/2 \text{ and} \tag{D-12}$$

$$\mathcal{E}(\mu_o^2) = \mathcal{E}(\mu_a^2) = 1/3; \tag{D-13}$$

now

$$\mu = \frac{1}{m} \sum_{1}^{m} \mu_o\ \mu_a \tag{D-14}$$

which is a sample estimate of $\mathcal{E}(\mu_o \mu_a)$, and has the mean value,

$$\bar{\mu} = \mathcal{E}(\mu_o \mu_a) \tag{D-15}$$

If the picture and aperture are uncorrelated, then

$$\bar{\mu} = \mathcal{E}(\mu_o \mu_a) = \mathcal{E}(\mu_o)\ \mathcal{E}(\mu_a) = 1/4 \tag{D-16}$$

On the other hand, if in some region, the two are perfectly correlated, i.e., $\mu_o = \mu_a$, then

$$\overline{\mu} = \mathcal{E}\{\mu_o \mu_a\} = \mathcal{E}\{\mu_o^2\} = 1/3 \quad . \tag{D-17}$$

By far the most common situation will be where the object and aperture are uncorrelated. Therefore, the average value of illuminance will be

$$\overline{E} = \frac{B}{64\pi} \cdot \frac{m}{n} \cdot \left(\frac{a_o}{a_o + a_i}\right)^2 \quad . \tag{D-18}$$

APPENDIX E

SURVEY OF METHODS FOR SCANNING PHOTOGRAPHIC DATA

1. Scanning System Requirements

A scanner is required that will accurately and rapidly convert photographic data into a video signal suitable for input to imagery screening circuitry. The full state-of-the-art capability must be utilized to obtain an optimum balance among (1) resolution, (2) speed of operation, (3) signal to noise ratio, (4) flexibility, and (5) cost.

The resolution capability required of a scanning device for aerial photograph imagery is dictated by several factors, among them:

a. size of film

b. resolution of film

c. ground resolution required

d. ground area to be covered, and

e. scanner limitations.

Giving due consideration to all of these factors, it has been concluded that a 5000 TV line capability is sufficient for imagery screening operations. Thus an element size of one foot may be obtained on a photograph covering 5000 feet by 5000 feet. An element size of one foot appears sufficiently small to permit recognition of military vehicles and other objects of comparable size.

a. Resolution

Scanning of 5000 elements in a single direction requires a highly refined scanning system in terms of resolution capability. It would be desirable to scan the 5000 elements in a single scan; however the possibility of duplicating scanning equipment for parallel operation may also be considered.

b. Speed of Operation

Rapid screening of photographic data imposes stringent speed requirements on the scanner. It is desirable to process an entire 25×10^6-element frame in a matter of a few seconds. Since ordinary scanners in current use operate below six megacycles, approximately eight seconds would be required to scan a data frame. It would be desirable to reduce this time requirement.

c. Signal-to-Noise Ratio

The signal-to-noise ratio requirement is dictated by the number of distinct levels to be discriminated. An adaptation of the channel capacity formula gives the relation between signal-to-noise ratio and number (L) of distinct levels, i.e.,

$$L = (1 + S/N)^{1/2} \tag{E-1}$$

where L is the number of levels, S is the mean signal power, and N is the noise power. For large signal-to-noise ratios, this equation reduces to

$$L = \frac{I_s}{I_n} \tag{E-2}$$

where I_s is rms signal current and I_n is rms noise current. Thus, the number of levels available is numerically equal to signal-to-noise ratio properly defined.

The availability of a given number of quantization levels determines the minimum signal-to-noise ratio that will justify use of all levels. Signal-to-noise ratio of scanning equipment is normally given in terms of peak current to the noise generated at that current, i.e.,

$$\left.\frac{I}{I_n}\right|_{I_{max}} = \frac{2\sqrt{2}\, I_s}{I_{n\,max}} \tag{E-3}$$

Since I_n (considered to be generated at a photosensitive surface) increases as the square root of signal current, its value at the mean signal level is $2^{-1/2}\, I_{n\,max}$, i.e., $I_{n\,max} = 2^{1/2}\, I_n$.

$$\left.\frac{I}{I_n}\right|_{I_{max}} = \frac{2\sqrt{2}\, I_s}{2^{1/2}\, I_n} = 2\,\frac{I_s}{I_n} \qquad \text{(E-4)}$$

$$\left.\frac{I}{I_n}\right|_{I_{max}} = 2\,L\,. \qquad \text{(E-5)}$$

Thus, for 64 level (6 bit) quantization, the indicated signal-to-noise ratio should be

$$\left.\frac{I}{I_n}\right|_{I_{max}} = 2 \times 64 = 128\,. \qquad \text{(E-6)}$$

A 128 to 1 (42 db) signal-to-noise ratio is necessary to justify six-bit quantization.

d. Scan Pattern Flexibility

Imagery screening applications may necessitate the use of area scan and other unusual patterns. The area scan pattern may take the form of a 60-element ribbon scan moving horizontally the full width of the scanner screen.

A 60-element ribbon height is required for registry testing of 30-element square areas. For one foot ground resolution element size, a 30-foot by 30-foot ground area is examined in the correlator; this permits recognition of military vehicles (typically 20 to 30 feet long) and objects of comparable size in a single parallel operation. Testing of all possible 30-element square arrays within a frame is facilitated by providing 50 percent vertical overlap, necessitating a 60-element high ribbon scan.

This overlap requirement results in unequal intervals between successive scans over any picture element; this is an important factor in the evaluation of storage type scanning devices. It should be mentioned that overlap scanning reduces the effective element rate capability of a scanning device.

e. Low Cost

While this program is directed toward future equipment development, it is desirable to determine capabilities on both the long and short term. For initial development equipment, certain sacrifices may be made to hold cost at a minimum. However, a view must be taken of future costs of more sophisticated equipment, and of possible cost reductions due to refinement of manufacturing techniques. More emphasis will be placed in this section on present costs, as they may be more accurately estimated.

2. Candidate Scanning Devices

a. Camera Tubes

There are two fundamental types of camera tubes available: storage types and non-storage types. Storage tubes depend upon the integration of photo-electronic effects during the time when no read-out is taking place. Non-storage types, on the other hand, utilize these effects only during the read-out interval. It should be noted that a constant storage interval is necessary in storage tubes, whereas a constant read-out interval is required in non-storage types.

Several camera tubes have been introduced since the advent of television, but only two or three of these have undergone continued development. The image orthicon has been refined to a high degree of reliability for use in television cameras, and the vidicon, because of its simplicity and small size, has been developed for industrial television purposes. Both of these fall into the storage tube category. The most interesting non-storage type is the image dissector, which is used for special purpose applications. A description of each of these tubes follows.

(1) Image Orthicon

The image orthicon combines an image intensifier section and a storage tube section. In addition, photomultiplication is accomplished in the same envelope. The combination results in a camera tube that has high sensitivity and an output current capable of overriding resistor and amplifier noise.

At the present, image orthicons are available with 900 TV line resolution. There are development programs under way to increase this to 1500 or more TV lines. There is some tendency toward burnt-in images on the image orthicon, but there is no theoretical lag to the signal. However, there seems to be little work in progress on utilizing the image orthicon for high frame rates.

(2) Vidicon

The vidicon is a simple photoconductive camera tube which can be made small in size while still retaining high resolution. The highest limiting resolution available in a vidicon is about 1200 TV lines. No known development is in progress for vidicons of higher resolution.

Vidicon frame rate is limited by the lag characteristic of the photoconductive target. The fastest decay time of present vidicons is about 0.1 second (for 10% retention); the data rate is limited by this lag characteristic.

(3) Image Dissector

This camera tube is of interest as an example of the non-storage type. Basically, it operates on the image intensifier principle. A photocathode emits electrons which are accelerated toward the rear of the tube. A pinhole aperture at the rear of the tube allows only a small stream of electrons to pass to the photomultiplier section. Deflection coils are used to shift the electron image to every possible condition of registry with respect to the aperture. Resolutions of about 3000 TV lines are available in these tubes, but adequate resolution, bandwidth, and signal-to-noise ratio cannot be obtained simultaneously. Image dissectors may be designed for particular applications, and are the nearest competitor in resolution capability to the flying spot scanner.

b. Flying Spot Scanner

A reliable device for scanning transparencies may be constructed by imaging a cathode ray tube raster onto the transparency and collecting the transmitted light in a photomultiplier. Alternatively, a line scan may be used, with mechanical shifting of the transparency to provide scan in the remaining dimension. High quality flying spot scanning requires a special tube with high resolution, high intensity, and low phosphor grain noise. A line scan tube can be constructed of higher quality than a raster scan tube as only one dimensional correction is required.

(1) Conventional Scanners

The conventional flying spot scanner utilizes a high resolution CRT operating at high intensity; this is necessary to overcome photomultiplier noise. A single transparency may be scanned repeatedly, or a new one

shifted into position after each scan. There is necessarily some time lost during the mechanical shifting operation. There are flying spot scanners that operate reliably with in excess of 2000 TV line resolution. The problem of mechanically shifting the transparency between frames is of such magnitude that an alternative should be sought.

(2) Line Scanner

Either a conventional scanner tube or a special purpose line scan tube may be used to generate a single line scan. With single line scan, continuous mechanical motion is necessary to provide vertical scanning. Continuous motion has the advantage of no loss in time between frames, and is simpler mechanically.

The chief disadvantage of line scan is that repeated scanning of the same line on the phosphor tends to cause heating, and possible damage to the phosphor or breakage of the tube. Thus, line brightness should not exceed the frame brightness limit given for frame scan.

(3) Rotating Phosphor Tube Scanner

In order to retain the advantage of line scan, and at the same time increase brightness capability, a rotating drum phosphor may be used. Continuous rotation of the drum permits cooling of one portion of the phosphor while another portion is being scanned. The brightness capability of such a system is about 30 times greater than with conventional line scan.

A rotating drum tube is available that is capable of a limiting resolution of up to 9000 TV lines, and which utilizes a fast P-16 phosphor.

APPENDIX F

FLYING SPOT SCANNER FOR HIGH SIGNAL-TO-NOISE RATIO AND HIGH RESOLUTION

1. Introduction

This appendix examines a flying spot scanner of the type required to scan large high resolution transparencies for imagery screening. The purpose of this examination is to determine whether desired resolution and signal-to-noise ratio performance may be realized in flying spot scanners of conventional design.

A functional diagram of the scanner considered is given in Figure F-1. The purpose of the auxiliary channel shown in this figure is to provide phosphor cancellation, i.e., to divide out any signal variations due to phosphor grain, phosphor fatigue, or other sources of non-uniformity prior to the first lens. Division is accomplished by logging the outputs of both channels and obtaining the difference. This difference is taken as the scanner output (instead of taking the inverse log).

2. Flying Spot Cathode Ray Tubes

Flying spot cathode ray tubes capable of 2500 TV line resolution on a 4-1/2-inch diameter screen are available from several manufacturers. These tubes are available with fast (P-16) phosphors to satisfy element rate requirements of up to 30×10^6 elements/second*.

Power input to a flying spot scanner may be taken as the product of cathode ray beam current and accelerating potential. Typical maximum beam current is about 10 microamperes for most high resolution tubes operating with reduced raster size; recommended accelerating potentials are between 8 and 27 kilovolts.

Useful light output from the cathode ray tube screen depends upon phosphor efficiency, which is typically 2 percent for P-16 operating at high resolution. This is the ratio of power emitted within the P-16 spectral response curve to total beam power. The spectral power distribution characteristic is an important consideration, since it must be utilized in determining lens transmission and photomultiplier response.

---

* Light output from a P-16 phosphor decays to 10 percent in about 0.1 microseconds, therefore, little degradation due to persistence is expected for video bandwidths below 5 megacycles.

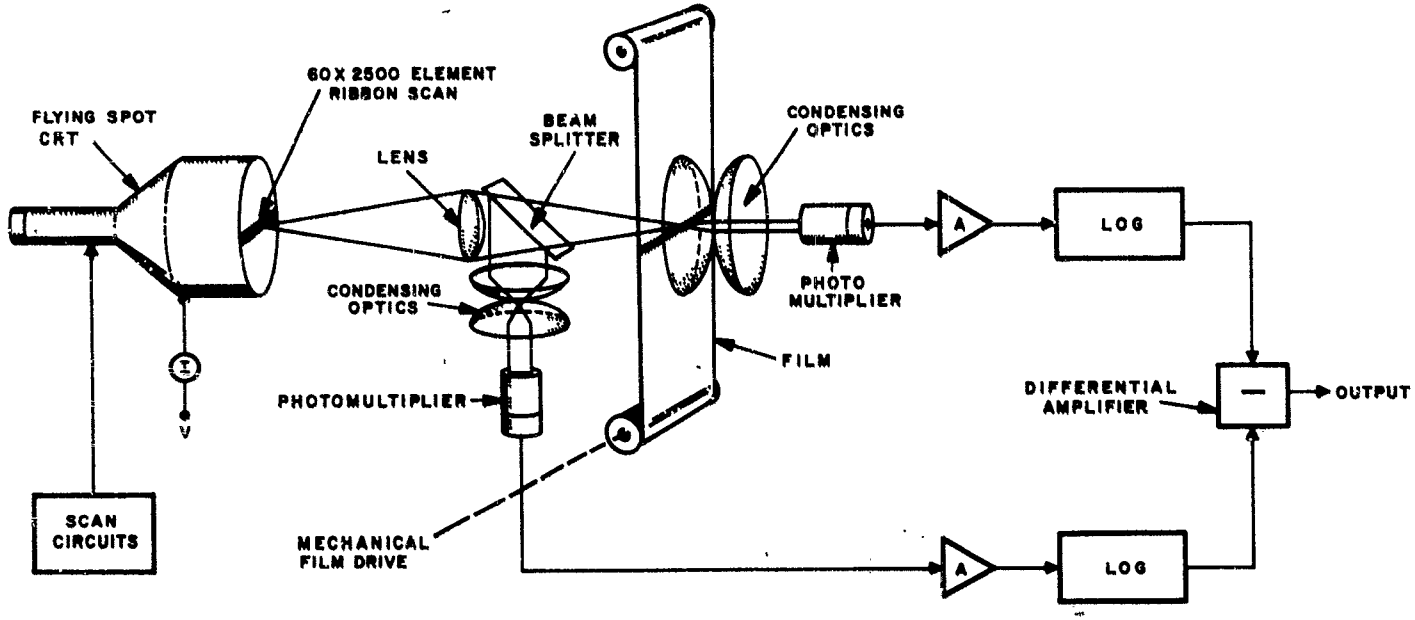


Figure F-1 Flying Spot Scanner With Phosphor Cancellation

3. Lens System

The lens system illustrated in the functional diagram (Figure F-1) comprises:

a. The objective lens

b. A beam splitter

c. Main channel condenser

d. Cancellation channel condenser

One-to-one magnification is assumed to result in a scanned region about 4-1/2-inches long on the transparency. Because image brightness is important in signal-to-noise ratio calculations, it is desirable to utilize an objective lens of as large f/number as possible; typical lenses for this application range from f/2 to f/8. While the resolution capability of the objective lens is an important consideration, it is assumed here that 5000 element resolution is available in an f/3.5 lens. Another important property of the objective is its spectral transmission characteristic. Since the P-16 phosphor peaks at the violet end of the visual range, a lens optimized in the visual range would have poor total transmission of the P-16 emission. For efficient transmission, therefore, it is necessary to utilize a lens optimized for the P-16 characteristic.

An objective will intercept only a small percentage of the light emitted from the CRT screen; the actual ratio of collected power to total power may be determined from the formula

$$\eta_c = \frac{1}{4f^2\ (m + 1)^2} \qquad \text{(F-1)}$$

where η_c is collection efficiency

f is the lens f number

m is magnification ratio.

For m = 1,

$$\eta_c = \frac{1}{16f^2} \qquad \text{(F-2)}$$

A beam splitter is used immediately behind the objective lens to direct some light into the cancellation channel. About 5 percent of the light should take this path, and the remainder should continue in its normal path to form a raster image on the transparency. Condensing optics are utilized to form objective aperture images on both photomultipliers. It is important that all optical components transmit efficiently over the P-16 phosphor spectrum. A 50 percent typical optical system loss may be assumed.

4. Photographic Data

Typical photographic transparencies have densities ranging between 0.2 and 2.0; corresponding transmission limits are about 0.60 to 0.01. It is difficult to obtain a good signal-to-noise ratio on high-density transparencies, therefore, it is advisable to ignore areas with transmission lower than 0.01, for example. This may be considered as the black level.

5. Photomultipliers

Linear photomultipliers (such as the "venetian blind" variety) are required for effective phosphor cancellation. The 8051 photomultiplier is a good example of this type: it has an S-11 spectral response that closely matches the P-16 phosphor spectrum, and has high cathode sensitivity of 0.06 amperes per watt at its peak. This tube also has reasonably uniform sensitivity over the photocathode surface, an important consideration when phosphor cancellation is necessary.

Photocathode sensitivity (η_k) to P-16 radiation is

$$\eta_k = \eta_p \frac{\int_\lambda f_1(\lambda)\, f_2(\lambda)\, d\lambda}{\int_\lambda f_1(\lambda)\, d\lambda} \qquad \text{(F-3)}$$

where: η_p = 0.06 amperes/watt = peak photomultiplier sensitivity

$f_1(\lambda)$ = P-16 spectral characteristic

$f_2(\lambda)$ = S-11 photocathode characteristic.

This calculation has been performed graphically, giving

$$\eta_k \approx 0.85\, \eta_p, \qquad \text{(F-4)}$$

indicating a very good spectral match.

The photocathode is the dominant source of noise in any flying spot scanner; in fact, useful computations may be made by ignoring all other noise sources. The shot noise formula

$$N_s = \sqrt{2ef\, i_k} \qquad \text{(F-5)}$$

is utilized to determine photocathode noise as a function of current. In Equation F-5

N_s = RMS noise current

e = charge of an electron = 16×10^{-20} coulombs

f = video bandwidth in cycles per second

i_k = photocathode current.

Photomultiplication raises the signal and noise levels linearly without introducing a significant amount of noise, and provides an output current sufficiently high to override resistor and tube noise in succeeding circuitry.

6. Logging

Logging accomplishes two functions in the scanner under consideration. It permits phosphor cancellation, and it introduces a non-linearity which causes output noise to remain more nearly constant as the signal output varies. This second advantage is more apparent when quantization of the output signal is necessary.

Phosphor decay compensation may be applied to both channels prior to logging; the two circuits should be linear and identical.

7. Differential Amplifier

This circuit merely subtracts cancellation channel output from main channel output to produce a signal free of phosphor grain noise and the other defects mentioned earlier.

8. Signal and Noise Considerations

Photomultiplier signal-to-noise ratio may be calculated easily as

$$\left(\frac{S}{N}\right)_{pm} = \frac{i_k}{\sqrt{2\, e\, f\, i_k}} . \qquad \text{(F-6)}$$

If very high signal-to-noise ratios are assumed, logging results in an output that is the logarithm of signal input plus noise of the same form as at the input but changed in magnitude (see Figure F-2). Noise amplitude is multiplied by the slope of the log characteristic. To illustrate this process, let the output current of a logarithmic device be represented as

$$i' = k \log i. \tag{F-7}$$

Then

$$n' \approx k\, n \frac{d \log i}{d i} = \frac{k}{i}\, n. \tag{F-8}$$

It is next required to determine what happens to signal and noise in the differential amplifier. Assume that subtraction is performed without loss or amplification. The inputs to the difference process are: main channel log signal plus noise (assumed Gaussian), and cancellation channel log signal plus noise (assumed Gaussian). The second is subtracted from the first to obtain:

$$i'_m - i'_c = k (\log i_m - \log i_c) = k \log (i_m/i_c) \tag{F-9}$$

and $$\sqrt{n'^2_m + n'^2_c} = \sqrt{(kn_m/i_m)^2 + (kn_c/i_c)^2}\,. \tag{F-10}$$

Note that the same log characteristic was utilized in main and cancellation channels.

9. Representative Performance Calculations

Figure F-3 represents in block form the operations that are performed in a flying spot scanner with phosphor grain cancellation. The functions follow closely those enumerated in the foregoing sections; exceptions are: (1) placement of the spectral match factor block immediately after the phosphor efficiency block; this makes separate computations for each photomultiplier unnecessary; and (2) neglect of light loss in the main channel due to beam splitting; this loss is lumped with the assumed lens system losses.

The following parameters are assumed for the representative system:

cathode-ray accelerating potential	27 KV
beam current	10 μa
phosphor efficiency	2 %

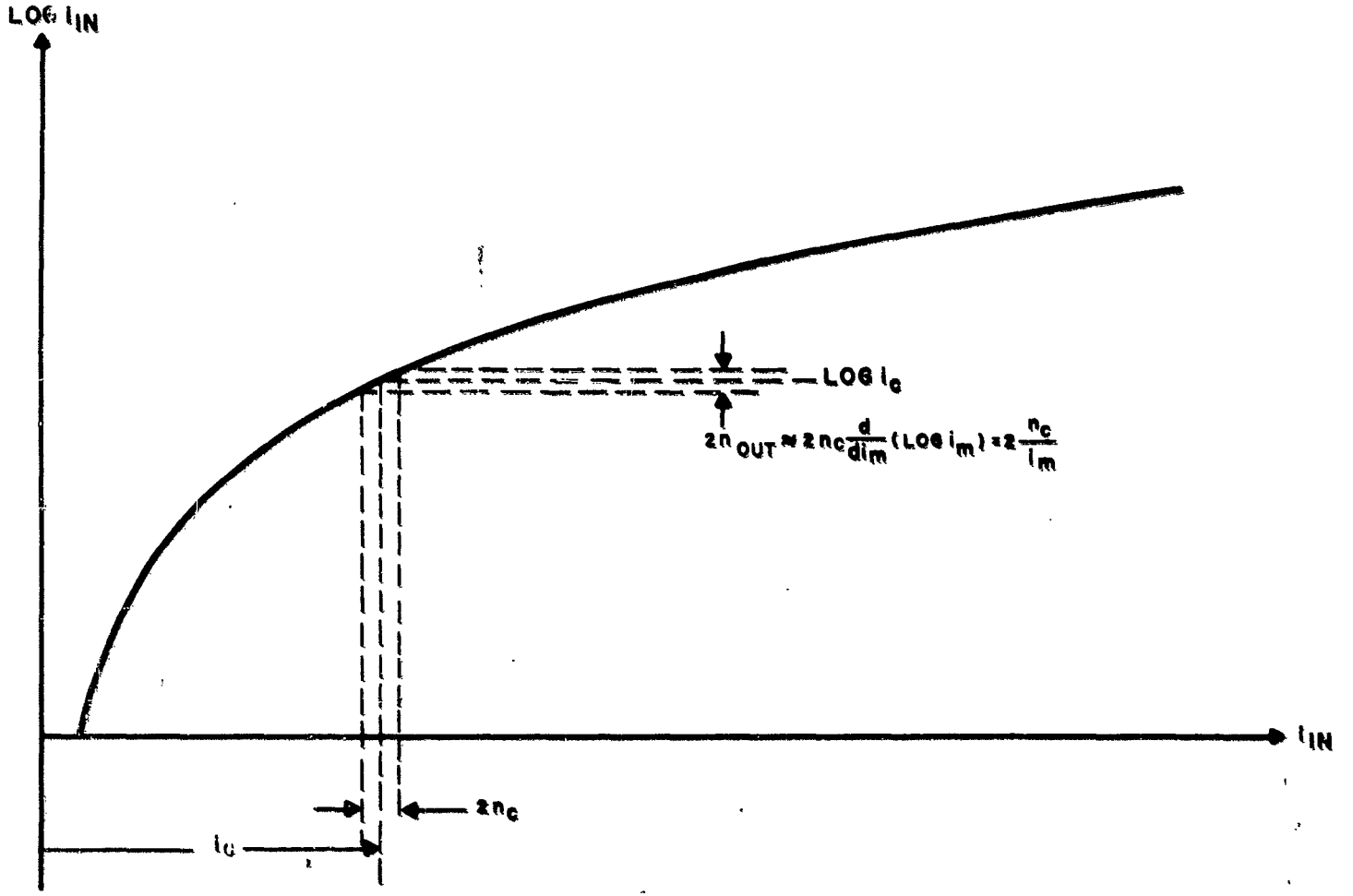


Figure F-2 Noise in Logarithmic Amplification

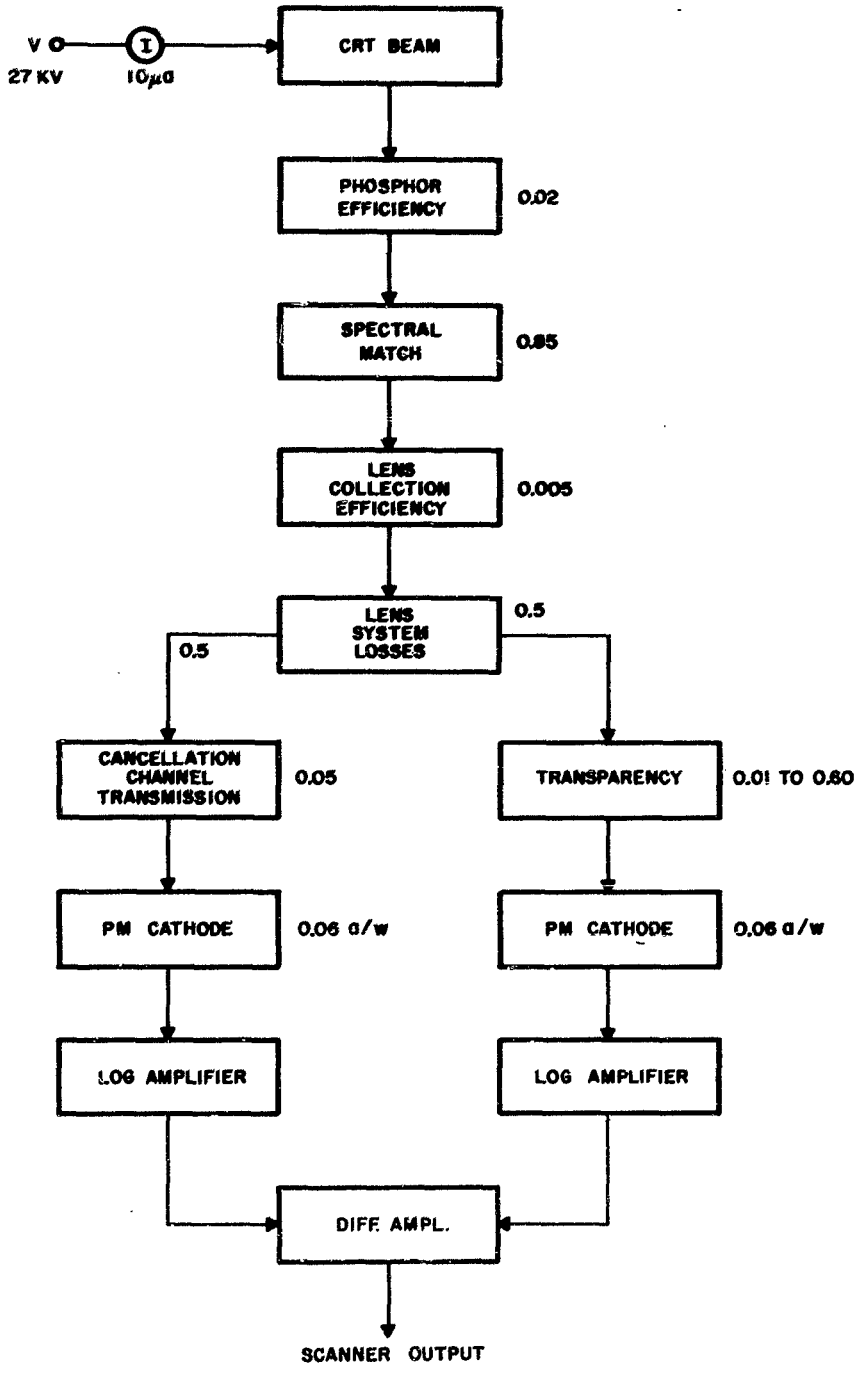


Figure F-3 Block Diagram of Essential Scanner Functions

spectral match (fixed)	85 %
objective collection efficiency (f/3.5 lens)	0.005
lens system losses	50 %
portion of light to cancellation channel	5 %
transmission range of film	1 % to 60 %
photocathode efficiency	0.06 a/w
video bandwidth	5 mc*

Utilizing these figures, cathode ray beam power is $27 \times 10^3 \times 10 \times 10^{-6} = 0.27$ watts; effective light output is two percent of this or 5.4 milliwatts. In terms of photomultiplier response, 85 percent of this power is usable, resulting in an effective radiant screen output of approximately 4.6 milliwatts. The amount of radiation intercepted by the objective lens is

$$W_o = \frac{4.6 \times 10^{-3}}{16\ (3.5)^2} = 23.5 \text{ microwatts.} \qquad \text{(F-11)}$$

Since 50 percent of this power is assumed to be lost in the optical system, available main channel radiant power to the film is 11.8 microwatts, and $W_m = 11.8\ \tau$ microwatts are delivered to the main channel photocathode (where τ is film transmission). Power delivered to the cancellation channel photomultiplier is five percent of 11.8 microwatts, i.e., $W_c = 0.590$ microwatts.

The photomultiplier conversion efficiency of 0.06 amperes per watt gives corresponding signal currents of

$$i_m = 0.06 \times 11.8\ \tau = 0.708\ \tau\ \mu a \qquad \text{(F-12)}$$

$$i_c = 0.06 \times 0.59 = 0.0354\ \mu a. \qquad \text{(F-13)}$$

---

* Phosphor decay compensation is assumed to have a negligible effect on performance calculations for a P-16 phosphor at 5 megacycles video bandwidth.

The shot noise formula (Equation F-5) gives RMS noise current at each photocathode as a function of signal current, i.e.,

$$n = \sqrt{2 \times 16 \times 10^{-20} \times 5 \times 10^{6}\ i}$$

$$= 1.265 \times 10^{-6} \sqrt{i} \text{ milliamperes.} \quad \text{(F-14)}$$

Thus, rms noise current may be determined for the main channel as a function of film transmission, and for the cancellation channel as a fixed value.

$$n_m = 1.265 \times 10^{-6} \sqrt{0.708 \times 10^{-6}\ \tau}$$

$$= 1.06 \times 10^{-9} \sqrt{\tau} \text{ amperes} \quad \text{(F-15)}$$

$$n_c = 1.06 \times 10^{-9} \sqrt{0.05}$$

$$= 2.37 \times 10^{-10} \text{ amperes.} \quad \text{(F-16)}$$

Main channel signal-tc-noise ratio on highlights (60 percent film transmission) may be determined at this point.

$$\left(\frac{S}{N}\right)_{pm} = \frac{0.7 \times 10^{-6} \times 0.6}{0.06 \times 10^{-9} \sqrt{0.6}} = 510 \text{ to } 1$$

$$= 54.1 \text{ decibels.} \quad \text{(F-17)}$$

This is the photomultiplier output signal-to-noise ratio without phosphor cancellation. We continue with the analysis to determine signal and noise characteristics after logging and differencing the two channels.

Note that the constant k appears as a factor in both Equations F-9 and F-10. This factor will assume a definite value in any actual scanner, depending upon the amount of linear amplification prior to logging, and upon current units selected. It is useful to determine the value of k that results in unity output signal excursion as film transmission varies from its minimum to maximum levels. The results so obtained are useful regardless of the amount of linear amplification and the log characteristic utilized. The following discussion assumes mathematical logging, and any logging device utilized should approximate this characteristic closely over its operating range.

To ascertain k for film transmission variations between one percent and 60 percent, we set

$$k \left[(\log i_{m\,max} - \log i_{m\,min}) - (\log i_{c\,max} - \log i_{c\,min})\right] = 1. \quad \text{(F-18}$$

But since i_c is invariant, it suffices to write

$$k \left[\log \frac{i_{m\ max}}{i_{m\ min}} \right] = 1, \tag{F-19}$$

or

$$k = \frac{1}{\log \left[\frac{i_{m\ max}}{i_{m\ min}} \right]} = \frac{1}{\log \left[\frac{\tau_{max}}{\tau_{min}} \right]} = 0.244 \tag{F-20}$$

Setting signal output to zero at $i_m = i_{m\ min}$, the expression for normalized signal output (S_{out}) is written as

$$S_{out} = 0.244 \left[\log \frac{i_m}{i_c} - \log \frac{i_{m\ min}}{i_c} \right]. \tag{F-21}$$

The expression for noise output (N_{out}) at the same i_m is

$$N_{out} = 0.244 \left[\left(\frac{n_m}{i_m} \right)^2 + \left(\frac{n_c}{i_c} \right)^2 \right]^{1/2}. \tag{F-22}$$

Equations F-21 and F-22 may be written in terms of the film transmission, τ, by substituting current equivalents from Equations F-12, F-13, F-15, and F-16. The resulting expressions for signal and noise output are

$$S_{out} = 0.244 \log 100\ \tau \tag{F-23}$$

and

$$N_{out} = 0.244 \times 10^3 \sqrt{\frac{2.24}{\tau} + 44.8}\ . \tag{F-24}$$

It is now a simple matter to plot equivalent output signal-to-noise ratio (R) in decibels as a function of τ:

$$R(\tau)\ \text{db} = 20 \log_{10} \left(\frac{1}{N_{out}} \right) \tag{F-25}$$

$$= 72.24 - 10 \log_{10} \left(\frac{2.24}{\tau} + 44.8 \right). \tag{F-26}$$

Plots of output signal as well as output signal-to-noise ratio as functions of τ are included in Figure F-4 for the scanner examined in this section. This figure is also very useful in determining signal-to-noise ratio for output signals smaller than the full range. Let $(S/N)_p$ represent signal-to-noise ratio in decibels for a signal covering P% of the output range. Then

$$\left(\frac{S}{N}\right)_p = \left(\frac{S}{N}\right)_{100} - 40 + 20 \log_{10} P \qquad \text{(F-27)}$$

For example, a signal covering 1 percent of the output range (about the smallest discernible change in a typical display) would result in signal-to-noise ratios 40 db less than those of Figure F-4.

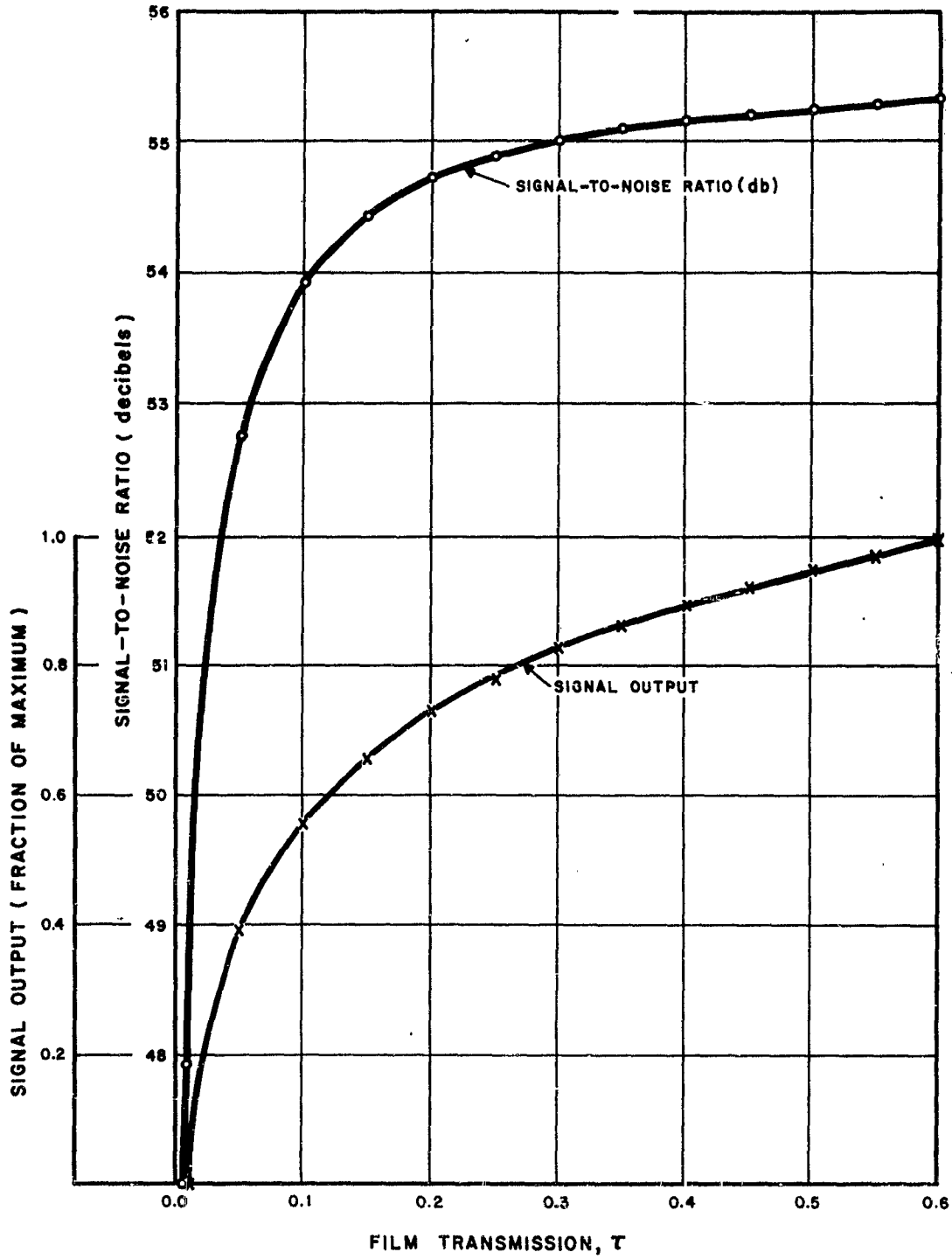


Figure F-4 Output and Signal-to-Noise Ratio of a Typical High Resolution Flying Spot Scanner for High Signal-to-Noise Performance

APPENDIX G

DETAILS OF GLASS DELAY LINE CORRELATOR

1. Bit Rates and Resolution

A typical glass delay line and transducer which is available from Corning, and which has been demonstrated to Philco, has a center frequency of 15 megacycles and a bandwidth of 7 to 10 megacycles. In systems which read out a true reproduction of the amplitude modulated carrier, an additional limiting factor on the resolution is imposed by the effect of the slit used for read-out. A single slit represents a different fraction of a wavelength for the higher frequency sidebands than for the carrier or the lower sidebands. Present single slit read-out uses a slit about one third of a carrier wavelength wide.

A 10 megabit data rate implies an ability to reproduce a 5 megacycle square wave with reasonable fidelity. If a video 3 db bandwidth of 5 mc is available, a train of binary input signals, and their corresponding propagated waves are represented in Figure G-1.

The 5 megacycle bandwidth restriction affects the ability of each bit representation to be independent of its predecessor. For example, the actual representation of zero is different in the eighth bit position in the sample train than it is in the second.

Several ways of improving the resolution seem possible. A single sideband driver might be designed to make better use of the available device bandwidth, or the pole pattern of the driving amplifier could be arranged to provide a wider overall characteristic.

A third possibility for improvement exists in the device itself. The curve of light transmission versus stress in the glass bar is nonlinear. In Figure G-1 the transfer characteristic is represented. Ordinarily, a bias plate is added in the optical path to place the operating point at 50% transmission where best linearity is obtained. In one of the proposed systems, however, the appropriate operating point is around zero transmission in order to produce rectification and therefore off-on transmission.

The effect of the biasing on the sample pulse train is represented in Figure G-1. The nonlinearity of the transfer characteristic has actually improved the ability to distinguish "0" from "1". If it were further possible to drive hard enough to enter the nonlinear portion of the curve at the high transmission end a further improvement could be obtained.

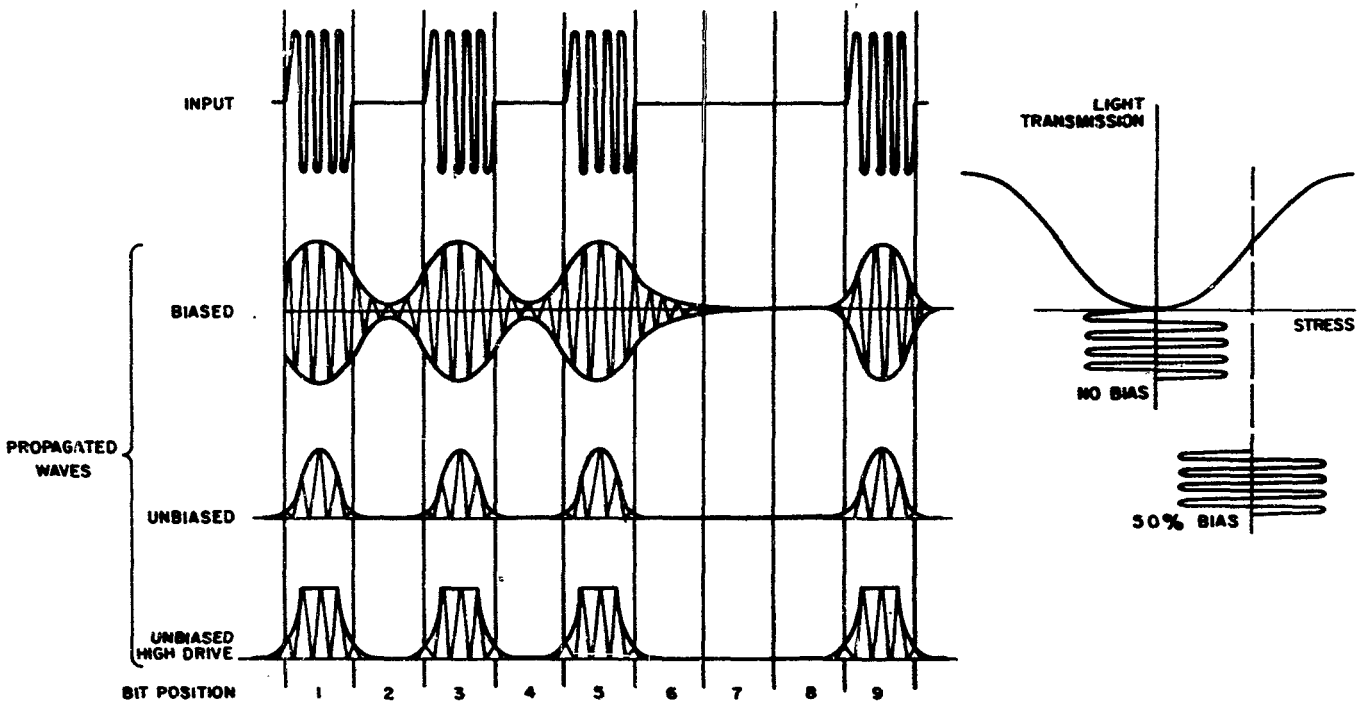


Figure G-1 Wave Propagation in Glass Delay Line

For a system utilizing slit read-out for retention of the carrier information, over drive of the line might provide a corresponding improvement. In fact, the bias plate might be more judiciously chosen to adjust the operating point to effect double clipping of the envelope information. The ability to double clip will also be improved by using a wider glass slab and higher driving power.

In summary,there is some improvement in definition to be had by special driving and operating techniques. Even without these however, it is possible to take the poor definition into account in the design of the mask sets by duplicating it in the input data used for mask evolution.

2. Drive Circuitry

The provision of adequate drive circuits for the transducer is a problem of some difficulty. About 20 watts of power must be delivered to a load of approximately 4 ohms in parallel with 4,000 pf at a carrier frequency of 15 megacycles. In addition, some bandwidth compensation should be attempted which will require even more power at the sidebands. A typical transducer bandwidth at 15 megacycles is 7 to 10 mc. Operating the transducer at a harmonic will not increase the bandwidth but may make it easier to design compensation circuits at the reduced bandwidth percentage.

For a feasibility study a vacuum tube driver provides the easiest solution, although a solid state driver developed by Wiley Electronics at a cost of $4800 can probably be obtained through Corning. It incorporates no compensation, however, and it seems more desirable to attempt a vacuum tube design for our purposes.

A high power pulsed oscillator * is available for initial measurements on the delay line.

3. Analog and Binary Light Modulation

Most of the experimental work with the optically tapped glass delay line has utilized linear operation. The optical set-up is as shown in Figure 4-6 The transfer characteristic is most linear at the 50% transmission point; the function of the quarter wave bias plate is to set operation around this point. A slit of less than approximately $\lambda/3$ is used to provide the read-out which is an accurate reproduction of the stress wave propogated in the glass, i.e., an amplitude modulated carrier. Recovery of the envelope alone then requires a normal RF detector.

---

* Arenberg PG650C - cost $1200

The choice of the 50% bias point is made for two reasons: this part of the transfer characteristic is most linear and this is also the point of maximum sensitivity.

For binary operation the same set-up can be used with subsequent detection of the RF signal. If the glass delay line is to be used for on-off light modulation, however, it becomes necessary to produce this rectification in the device. The operating point can be shifted to the non-linear portions of the curve.

4. Photosensor Collection Systems

A practical problem exists in collecting the light spread over a 15" length in order to properly illuminate a 1/2" photomultiplier input. The neatest solution is undoubtedly a fiber optics assembly to provide the necessary change and the mechanical freedom necessary for physically arranging 100 viewing photomultipliers (one per class). Each unit of the assembly would consist of a 15" long array, a fraction of an inch high, composed of uniformly distributed fibers at one end which are gathered into a round or square cross section bundle at the other end.

Corning utilized an array of 6 units, 0.6" high by 30 mils wide ending in circular cross sections for simultaneous 6-slit readout. The assembly was furnished by Bausch and Lomb for about $200. This gives some indication of cost and feasibility for the kind of assembly under discussion.

5. Homogeneity of Delay Line

Discussion with Corning's engineers indicates that good uniformity in the delay line can be expected throughout the individual unit. Also, the temperature coefficient of expansion is small enough (80 ppm) to not cause difficulty. Some problem may be experienced in matching one delay line against another. Plots of read-out signal amplitude over the length of the line shows less than 3 db variation (using slit read-out) and verify the uniformity of the line. The uniformit on the output face in the direction normal to the length is not as good, however, due to the radiation pattern of the transducer. In addition, since the transducer covers only about 80% of the cross section of the glass, the edges of the line are not usable for read-out. A plot of read-out photomultiplier current shows a variation of about $\pm$ 30 to $\pm$ 50% across the usable width. It may be necessary to provide a compensating mask or to arrange the optical system in such a way that each mask pair views the whole delay line face. There appears to be sufficient symmetry to count on a division of the face along the centerline with the top half for positive masks and the bottom half for the negative masks if the system requires it.

APPENDIX H

STATISTICAL METHODS FOR PATTERN CLASSIFICATION I

To make this report as self-contained as possible, this appendix presents a number of items related to the tutorial discussion on statistical methods for pattern classification.

1. The Likelihood Ratio

For the case of two groups, when the distributions of the populations from which samples are obtained are completely specified, Welch* showed in 1939 that a general discriminant function is the likelihood ratio of the two hypotheses. Suppose a number of measurements are made on a pattern and, on the basis of the measurements,

$$x = (x_1, x_2, \ldots, x_N),$$

it is desired to classify the pattern into one of the two groups to which it can possibly belong. We can think of the universe of patterns as an N-dimensional space and the task of classification as one of dividing this N-dimensional space into two mutually exclusive regions, R_1 and R_2, such that when a particular measurement x falls in R_1, the pattern is listed under Group 1 and when x falls in R_2, the pattern is listed under Group 2.

Let the proportion of the two groups in the universe from which the patterns are obtained be $q_1:q_2$, where $q_1 + q_2 = 1$. If $f_1(x)$ and $f_2(x)$ represent the probability density functions for the two populations, then the probability of making an error in the classification of patterns into the two groups is

$$q_1 \int_{R_2} f_1(x)\, dx + q_2 \int_{R_1} f_2(x)\, dx, \tag{H-1}$$

where dx is the element of volume $(dx_1 \cdot dx_2 \cdot \ldots dx_N)$. Rewriting the expression in H-1 as

$$q_2 + \int_{R_2} (q_1 f_1 - q_2 f_2)\, dx \tag{H-2}$$

---

* Welch, B. L., "Note on Discriminant Functions," Biometrika, Vol. 31, pp. 218-220, 1939.

it is easy to see that the probability of misclassification is minimized by putting all those points in the decision space for which $q_1 f_1 < q_2 f_2$ into the Region R_2. Thus the regions which give the minimum value for the probability of misclassification are defined in terms of the likelihood ratio $(f_1)/(f_2)$.

$$R_1 : \frac{f_1}{f_2} \geqslant \frac{q_2}{q_1}$$

$$R_2 : \frac{f_1}{f_2} < \frac{q_2}{q_1} \qquad \text{(H-3)}$$

From H-3 we see that when q_1 and q_2 are known and fixed, the regions R_1 and R_2 are separated by a boundary along which the likelihood ratio has a constant value. Here points on the boundary have been arbitrarily assigned to R_1.

2. Some Decision Criteria

Situations in which different misclassifications have different consequences may arise. Such situations are formulated in terms of Wald's theory of decision functions* which he presented in 1939 for finite and infinite alternative. In 1945, Von Mises** gave the solution which minimizes the maximum error of classification for the case of finite alternatives. Further results and applications of decision theory to classification problems were given in a series of papers by Rao.***

Let C_{21} be the loss incurred when a pattern from Group 1 is assigned to Group 2 and C_{12} the loss resulting in assigning a pattern from Group 2 to Group 1. Then the expected loss is

$$q_1 C_{21} \int_{R_2} f_1 \cdot dx + q_2 C_{12} \int_{R_1} f_2 \, dx \qquad \text{(H-4)}$$

---

* Wald, A., "Contributions to the Theory of Statistical Estimation and Testing Hypotheses," Ann. Math. Stat., Vol. 10, pp. 299-326, 1939.

** Von Mises, R., "On the Classification of Observation Data into Distinct Groups," Ann. Math. Stat., Vol. 16, pp. 68-73, 1945.

*** Rao, C. R., "Statistical Inference Applied to Classificatory Problems," Sankhya, Vol. 10, pp. 229-256, 1950; Vol. 11, pp. 107-116, 1951; Vol. 12, pp. 229-246, 1952-1953.

and the solution which minimizes this is

$$\left.\begin{aligned} R_1 : \quad & \frac{f_1}{f_2} \geqslant \frac{q_2\, C_{12}}{q_1\, C_{21}} \\ R_2 : \quad & \frac{f_1}{f_2} < \frac{q_2\, C_{12}}{q_1\, C_{21}} \end{aligned}\right\} \qquad \text{(H-5)}$$

Note that only the value of the threshold t, against which the likelihood ratio is compared, has changed. The above value of the threshold resulted from using Bayes criterion, viz., minimization of the expected loss of classification. Here we had tacitly assumed that no gain resulted from a correct decision; the general result using a Bayes criterion gives

$$t = \frac{q_2\,(C_{12} - C_{22})}{q_1\,(C_{21} - C_{11})} . \qquad \text{(H-6)}$$

For a priori probabilities and costs for which t becomes 1 we obtain the maximum - likelihood criterion. When a priori probabilities are not known, the principle of maximum likelihood leads to a Bayes procedure with equal a priori probabilities assigned to each alternative. Another possibility is the minimax decision rule* which is the Bayes rule relative to the a priori distribution for which the expected loss is a maximum. In all these cases the decision rule is

$$\left.\begin{aligned} R_1 : \quad & \frac{f_1}{f_2} \geqslant t \\ R_2 : \quad & \frac{f_1}{f_2} < t \end{aligned}\right\} \qquad \text{(H-7)}$$

3. Likelihood Ratio for Multivariate Normal Distributions

The joint probability density function of N gaussian real random variables x_i with mean m_i and variances v_{ii} is

* Anderson, T. W., "An Introduction to Multivariate Statistical Analysis," New York, John Wiley, 1958. Chapter 6.

$$f(x_1, x_2, \ldots, x_N) =$$

$$\frac{1}{(2\pi)^{N/2} |V|^{1/2}} \exp \left[-\frac{1}{2|V|} \sum_{i=1}^{N} \sum_{j=1}^{N} |V|_{ij} (x_i - m_i)(x_j - m_j) \right],$$

where $|V|_{ij}$ is the cofactor of the elements v_{ij} in the determinant $|V|$ of the covariance matrix

$$V = \begin{bmatrix} v_{11} & v_{12}, \ldots, v_{1N} \\ v_{21} & v_{22}, \ldots, v_{2N} \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ v_{N1} & v_{N2}, \ldots, v_{NN} \end{bmatrix}$$

in which

$$v_{ij} = E\left[(x_i - m_i)\ (x_j - m_j) \right] .$$

Denoting the vector of means by

$$M = (m_1, m_2, \ldots, m_N),$$

the joint probability density in matrix notation becomes

$$f(x) = \frac{1}{(2\pi)^{N/2} |V|^{1/2}} \exp \left[-\frac{1}{2} (X - M)' V^{-1} (X - M) \right] .$$

Now if N measurements are made on a pattern which can belong to one of two groups in which (a) the measurements are normally distributed – no assumptio

of statistical independence of measurements is made and (b) the covariance matrices for the two groups are equal, then the likelihood ratio is the ratio of two multivariate normal density functions which differ only in their mean vectors:

$$\frac{f_1(x)}{f_2(x)} = \frac{\exp\left\{-\frac{1}{2}\,(x - M^{(1)})'\,V^{-1}\,(x - M^{(1)})\right\}}{\exp\left\{-\frac{1}{2}\,(x - M^{(2)})'\,V^{-1}\,(x - M^{(2)})\right\}}$$

$$= \exp\left\{-\frac{1}{2}\left[(x - M^{(1)})'\,V^{-1}\,(x - M^{(1)}) - (x - M^{(2)})'\,V^{-1}\,(x - M^{(2)})\right]\right\} .$$

The condition of H-7 may now be applied to determine the set of x's which should be classified into Group 1 and Group 2, respectively. Taking the logarithm of the last expression gives, for the boundary between R_1 and R_2:

$$-\frac{1}{2}\left[(x - M^{(1)})'\,V^{-1}\,(x - M^{(1)} - (x - M^{(2)})'\,V^{-1}\,(x - M^{(2)})\right] = \log t,$$

which, after rearranging terms, becomes

$$x'\,V^{-1}\,(M^{(1)} - M^{(2)}) = \log t + \frac{1}{2}\,(M^{(1)} + M^{(2)})'\,V^{-1}\,(M^{(1)} - M^{(2)}). \quad \text{(H-8)}$$

The terms on the right-hand side of the above expression can be represented by a constant C, so that the regions R_1 and R_2 in the N-dimensional space are separated by a hyperplane whose equation is

$$x'\,V^{-1}\,(M^{(1)} - M^{(2)}) = C, \quad \text{(H-9)}$$

or

$$\sum_{i=1}^{N}\left(v^{1i}d_1 + v^{2i}d_2 + \ldots + v^{Ni}d_N\right)x_i = C \quad \text{(H-10)}$$

where

v^{ji}, (i, j = 1, 2, . . . N) are elements of V^{-1}

and

$$d_j = (m_j^{(1)} - m_j^{(2)}).$$

When the covariance matrices for the two populations are not assumed to be equal, the likelihood ratio is

$$\frac{f_1(x)}{f_2(x)} = \frac{|V_2|^{1/2}}{|V_1|^{1/2}} \cdot \frac{\exp\left\{ -\frac{1}{2}(x - M^{(1)})' V_1^{-1} (x - M^{(1)}) \right\}}{\exp\left\{ -\frac{1}{2}(x - M^{(2)})' V_2^{-1} (x - M^{(2)}) \right\}}$$

The logarithm of the likelihood ratio, after rearranging terms, is

$$-\frac{1}{2} x' \left[V_1^{-1} - V_2^{-1} \right] x + x' \left[V_1^{-1} M^{(1)} - V_2^{-1} M^{(2)} \right]$$

$$-\frac{1}{2} M^{(1)'} V_1^{-1} M^{(1)} + \frac{1}{2} M^{(2)'} V_2^{-1} M^{(2)} + \log \left(\frac{|V_2|}{|V_1|} \right)^{1/2}.$$

The surface of constant likelihood ratio is then defined by

$$-\frac{1}{2} x' \left[V_1^{-1} - V_2^{-1} \right] x + x' \left[V_1^{-1} M^{(1)} - V_2^{-1} M^{(2)} \right] = \text{constant}, \qquad \text{(H-11)}$$

or

$$\sum_{i=1}^{N} \sum_{j=1}^{N} \left[v_1^{ij} (x_i - m_i^{(1)}) (x_j - m_j^{(1)}) - v_2^{ij} (x_i - m_i^{(2)}) (x_j - m_j^{(2)}) \right]$$

$$= \text{constant},$$

where v_1^{ij} and v_2^{ij} are elements of V_1^{-1} and V_2^{-1}, respectively. The last two expressions may be written in the form:

$$\sum_{\substack{i=1\\j=1}}^{N} a_{ij} x_i x_j + \sum_{i=1}^{N} b_i x_i = \text{constant}, \tag{H-12}$$

where

$$a_{ij} = \left(v_1^{ij} - v_2^{ij} \right) \tag{H-13}$$

and

$$b_i = 2\left[\left(m_1^{(1)} v_1^{i1} - m_1^{(2)} v_2^{i1} \right) + \left(m_2^{(1)} v_1^{i2} - m_2^{(2)} v_2^{i2} \right) \right.$$

$$\left. + \ldots + \left(m_N^{(1)} v_1^{iN} - m_N^{(2)} v_2^{iN} \right) \right] . \tag{H-14}$$

4. Likelihood Ratio when the Variables are Binary

Let $X = (X_1, X_2, \ldots, X_N)$, $X_i = 0$ or 1 for $i = 1, 2, \ldots, N$, denote the observables and let $f_1(x)$ and $f_2(x)$ denote the joint probability functions of the X_i in Group 1 and Group 2, respectively. In order to compute the likelihood ratio, it is necessary to consider expansions for the probability functions. Any parametric representation of an arbitrary distribution of N binary variables will in general require $(2^N - 1)$ independent parameters. A useful representation is an orthogonal expansion for joint probability functions of binary variables as in Bahadur*,**. Define the following parameters for the two groups ($g = 1, 2$; $i = 1, 2, \ldots, N$):

* Bahadur, R.R., "A Representation of the Joint Distribution of Responses to n Dichotomous Items," USAF SAM series in statistics, Report No. 59-42, Randolph AFB, Texas, 1959; appears in Studies in Item Analysis and Prediction Ed. Herbert Solomon, Stanford University Press, Stanford, California, 1961.

** Bahadur, R.R., "On Classification Based on Responses to n Dichotomous Items," USAF SAM series in Statistics, Randolph AFB, Texas, 1959; appears in Studies in Item Analysis and Prediction, Ed. Herbert Solomon, Stanford University Press, Stanford, California, 1961.

$$
\left.
\begin{aligned}
m_i^{(g)} &= \text{prob. } (x_i^{(g)} = 1); \quad 0 < m_i^{(g)} < 1 \\
y_i^{(g)} &= \frac{(x_i - m_i^{(g)})}{\sqrt{m_i^{(g)} \ (1 - m_i^{(g)})}} \\
r_{ij}^{(g)} &= E_g \ (y_i^{(g)} \ y_j^{(g)}) , \quad i < j \\
r_{ijk}^{(g)} &= E_g \ (y_i^{(g)} \ y_j^{(g)} \ y_k^{(g)}) ; \quad i < j < k \\
r_{1,2,\ldots N}^{(g)} &= E_g \ (y_1^{(g)} \ y_2^{(g)} , \ldots \ y_N^{(g)})
\end{aligned}
\right\} \qquad \text{(H-15)}
$$

where E_g denotes that the expectation is taken with respect to the probability function of Group g, g = 1, 2. The $r_{ij}^{(1)}$, $r_{ijk}^{(1)}$, . . . , and $r_{ij}^{(2)}$, $r_{ijk}^{(2)}$, . . . are the "correlation parameters" for Group 1 and Group 2, respectively. Let $p_1^{(g)}(x)$ denote the probability functions of the X_i in the two groups when the X_i are independently distributed but have the same mean values $m_i^{(g)}$, as they do when $f_g(x)$ hold. Then

$$
p_1^{(g)}(x) = \prod_{i=1}^{N} \left[m_i^{(g)}\right]^{x_i} \left[1 - m_i^{(g)}\right]^{(1 - x_i)} \qquad \text{(H-16)}
$$

The functions $\dfrac{f_g(x)}{p_1^{(g)}(x)}$ can be expanded in the orthogonal expansions given by:

$$\frac{f_g(x)}{p_1^{(g)}(x)} = 1 + \sum_{i<j} r_{ij}^{(g)}\, y_i^{(g)}\, y_i^{(g)} + \sum_{i<j<k} r_{ijK}^{(g)} y_i^{(g)}\, y_j^{(g)}\, y_k^{(g)} + \ldots + r_{12,\ldots N}^{(g)}\, y_1^{(g)}\, y_2^{(g)} \ldots, y_N^{(g)} \qquad \text{(H-17)}$$

Denoting the right side of Equation H-17 by $h^{(g)}(x)$,

$$f_g(x) = p_1^{(g)}(x) \cdot h^{(g)}(x) \qquad \text{(H-18)}$$

The likelihood ratio $f_1(x)/f_2(x)$ is then given by

$$L(x) = \frac{p_1^{(1)}(x)\; h^{(1)}(x)}{p_1^{(2)}(x)\; h^{(2)}(x)}$$

or

$$L(x) = \frac{\left\{\prod_{i=1}^{N}\left[m_i^{(1)}\right]^{x_i}\left[1-m_i^{(1)}\right]^{(1-x_i)}\right\}\left\{1 + \sum_{i<j} r_{ij}^{(1)} y_i^{(1)}\, y_i^{(1)} + \ldots\right\}}{\left\{\prod_{i=1}^{N}\left[m_i^{(2)}\right]^{x_i}\left[1-m_i^{(2)}\right]^{(1-x_i)}\right\}\left\{1 + \sum_{i<j} r_{ij}^{(2)} y_i^{(2)}\, y_i^{(2)} \ldots\right\}} \qquad \text{(H-19)}$$

By dropping terms of $h^{(g)}(x)$, different approximations can be obtained for $f_g(x)$ from Equation H-18, and the result will be a legitimate probability function as long as the terms of $h^{(g)}(x)$ which are retained are non-negative for all x. The first-order approximation to $f_g(x)$ is $p_1^{(g)}(x)$. The second-order approximation is

$$p_2^{(g)}(x) = p_1^{(g)}(x) \cdot \left[1 + \sum_{i<j} r_{ij}^{(g)}\, y_i^{(g)}\, y_j^{(g)}\right], \qquad \text{(H-20)}$$

and so on. When first-order approximations to $f_1(x)$ and $f_2(x)$ are used,

the logarithm of the likelihood ratio is easily seen to be

$$\text{Log } L(x) = \sum_{i=1}^{N} (a_i x_i + c_i), \tag{H-21}$$

where

$$a_i = \log \left[\frac{m_i^{(1)} (1-m_i^{(2)})}{m_i^{(2)} (1-m_i^{(1)})} \right] \tag{H-22}$$

and

$$c_i = \log \frac{(1 - m_i^{(1)})}{(1 - m_i^{(2)})} \tag{H-23}$$

When the second-order approximation of Equation H-20 is used for $f_1(x)$ and $f_2(x)$, the logarithm of the likelihood ratio is

$$\text{Log } L(x) = \sum_{i=1}^{N} (a_i x_i + c_i) + \log \left[1 + \sum_{i<j} r_{ij}^{(1)} y_i^{(1)} y_j^{(1)} \right] - \log \left[1 + \sum_{i<j} r_{ij}^{(2)} y_i^{(2)} y_j^{(2)} \right] \tag{H-24}$$

where a_i and c_i are as defined in Equations H-22 and H-23 and the other terms are as defined in Equation H-15.

If a digital computer is being used for computation, the expression in Equation H-24 poses no major problem. If the "correlation parameters" are small enough so that one can use the approximation $\log (1+\theta) \approx \theta$, then a function which is somewhat easier to implement than that of Equation H-24 is obtained, viz.,

$$\text{Log } L(x) \approx \sum_{i=1}^{N} \left[a_i + \sum_{j \neq r}^{N} (-m_j^{(1)} u_{ij}^{(1)} + m_j^{(2)} u_{ij}^{(2)}) \right] x_i + \sum_{i<j} (u_{ij}^{(1)} - u_{ij}^{(2)}) x_i x_j + \sum_{i=1}^{N} c_i, \tag{H-25}$$

where

$$u_{ij}^{(g)} = \frac{r_{ij}^{(g)}}{\sqrt{m_i(g)\ (1-m_i(g))\ m_j(g)\ (1-m_j(g))}}, \quad g = 1,2.$$

5. The Anderson-Bahadur Plane Boundary

The recent paper by Anderson and Bahadur* presents the class of best (admissible) linear procedures for the case of two multivariate normal distributions which differ in mean vectors and have unequal covariance matrices. For the case of arbitrary distributions, i.e., distributions not necessarily normal, the procedure finds a direction of projection for the two samples and computes a linear function for which the ratio of the difference between the means of the projected samples to the sum of the standard deviations of the projected samples is maximized. If the direction of projection is given by the column vector $A = (a_1, a_2, \ldots, a_N)$ and $M^{(1)}$ and $M^{(2)}$ are the mean vectors for Group 1 and Group 2, with V_1 and V_2 as the corresponding covariance matrices, the minimax Anderson-Bahadur procedure finds the vector A which maximizes the ratio

$$\frac{A' (M^{(1)} - M^{(2)})}{(A' V_1 A)^{1/2} + (A' V_2 A)^{1/2}} \qquad \text{(H-26)}$$

The resulting A is given by

$$A = \left[t V_1 + (1-t) V_2 \right]^{-1} (M^{(1)} - M^{(2)}), \qquad \text{(H-27)}$$

where t is a scalar $0 < t < 1$ given by

$$A' \left[t^2 V_1 + (1-t^2) V_2 \right] A = 0. \qquad \text{(H-28)}$$

Starting with an initial value for t, A is obtained from Equation H-27; these values for t and A are then used to see if Equation H-28 is satisfied. In this manner A is solved for iteratively. In terms of the vector of measurements $X = (x_1, x_2, \ldots, X_N)$, the plane boundary for classification is given by

$$A'X = C, \qquad \text{(H-29)}$$

*Anderson, T.W. and Bahadur, R.R., "Classification Into Two Multivariate Normal Distributions with Different Covariance Matrices," Ann. Math. Stat., Vol. 33, No. 2, pp. 42-431, June 1962; presented at I.M.S. Meeting, January 1960.

where

$$C = \frac{(A' V_2 A)^{1/2} \cdot A'M^{(1)} + (A'V_1 A)^{1/2} A'M^{(2)}}{(A'V_1 A)^{1/2} + (A'V_2 A)^{1/2}} . \qquad \text{(H-30)}$$

In computing the above expressions when we have n_1 samples from Group 1 and n_2 samples from Group 2, the quantities $M^{(1)}$ and $M^{(2)}$ would be replaced by $\bar{x}^{(1)}$ and $\bar{x}^{(2)}$, respectively, and V_1 and V_2 by S_1 and S_2, respectively, where

$$\bar{x}(g) = (\bar{x}_1(g), \bar{x}_2(g), \ldots, \bar{x}_N(g)), \quad g = 1, 2, \qquad \text{(H-31)}$$

$$\bar{x}_i(g) = \frac{1}{n_g} \sum_{r=1}^{n_g} x_{i\,r}^{(g)}, \quad i = 1, 2, \ldots, N; \; g = 1, 2, \qquad \text{(H-32)}$$

and

$$S_g = \left[s_{ij}^{(g)} \right] \qquad \text{(H-33)}$$

where

$$s_{ij}^{(g)} = \frac{1}{(n_g - 1)} \sum_{r=1}^{n_g} (x_{i\,r}^{(g)} - \bar{x}_i^{(g)}) (x_{j\,r}^{(g)} - \bar{x}_j^{(g)}) . \qquad \text{(H-34)}$$

6. Multidimensional Scatter*, **

Let us rewrite the expression of Equation H-34, for the elements of the sample covariance matrix, as follows:

$$(n_g - 1) \; s_{ij}^{(g)} = u_{ij}^{(g)} = \sum_{r=1}^{n_g} (x_{i\,r}^{(g)} - \bar{x}_i^{(g)}) (x_{j\,r}^{(g)} - \bar{x}_j^{(g)}) . \qquad \text{(H-35)}$$

---

* Wilks, S.S., "Multidimensional Statistical Scatter," Contributions to Probability and Statistics in Honor of Harold Hottelling, Stanford Univ. Press, 1960.

** Wilks, S.S., Mathematical Statistics, New York, John Wiley and Sons, 1962.

Wilks uses the name internal scatter matix for the matrix $\left[u_{ij}^{(g)}\right]$ and he calls the determinant $\left|u_{ij}^{(g)}\right|$ the internal scatter of the sample g, g = 1, 2. It is instructive to consider the origin of this terminology.

Let

$$x_r = (x_{1r}, x_{2r}, \ldots x_{Nr}), \quad r = 1, 2, \ldots, n$$

be a sample of size n from an N- dimensional distribution, and let $(x_{10}, x_{20}, \ldots, x_{N0})$ be some pivotal point about which we wish to define the scatter of the sample. Define the matrix

$$H = \begin{bmatrix} (x_{11} - x_{10}), \ldots, (x_{N1} - x_{N0}) \\ \vdots \qquad\qquad \vdots \\ (x_{1n} - x_{10}), \ldots, (x_{Nn} - x_{N0}) \end{bmatrix} \tag{H-36}$$

Then the scatter matrix is

$$H'H = \begin{bmatrix} \sum_{r=1}^{n} (x_{1r} - x_{10})^2 & \ldots, & \sum_{r=1}^{n} (x_{1r} - x_{10})(x_{Nr} - x_{N0}) \\ \vdots & & \vdots \\ \sum_{r=1}^{n} (x_{Nr} - x_{N0})(x_{1r} - x_{10}), & \ldots, & \sum_{r=1}^{n} (x_{Nr} - x_{N0}) \end{bmatrix} \tag{H-37}$$

The scatter denoted by ${}_N S_{x_0, n}$ is defined as the determinant of H'H, i.e.,

$${}_N S_{x_0, n} = \left|H' H\right| . \tag{H-38}$$

The geometrical interpretation of the scatter is as follows. The sample of size n and the pivotal point can be represented as n+1 points in an N-dimensional Euclidean space. Consider the N+1 points obtained by using N out of the n sample points and the pivotal point. There are N+1 ways of choosing an additional point which, together with the N+1 points already picked, will form an N-dimensional parallelotope. Note that N+1 different parallelotopes can result but they will all have the same absolute value for their N-dimensional volumes. This absolute value is

called the N-dimensional content determined by $(x_{1r}, x_{2r}, \ldots, x_{Nr})$, $r = 1, 2, \ldots, n$ and $(x_{10}, x_{20}, \ldots, x_{N0})$. Now there are $\binom{n}{N}$ different ways of choosing N out of n sample points. The scatter defined in Equation H-38 is equal to the sum of the squares of the N-dimensional contents determined by $(x_{10}, \ldots, x_{N0})$ and each of the $\binom{n}{N}$ different possible choices of $(x_{1r}, x_{2r}, \ldots, X_{Nr})$, $r = 1, 2, \ldots, n$.

Now,

$$\sum_{r=1}^{n}(x_{ir}-x_{io})(x_{jr}-x_{jo}) = \sum_{r=1}^{n}(x_{ir}-\bar{x}_i)(x_{jr}-\bar{x}_j) + n(x_i-x_{io})(x_{jr}-x_{jo})$$

$$= u_{ij} + n(\bar{x}_i-x_{io})(\bar{x}_j-x_{jo}). \qquad \text{(H-39)}$$

Thus Equation H-38 can be written as

$$N^{S}_{x_o,\,n} = \begin{vmatrix} u_{11}+n(\bar{x}_1-x_{10})^2, \ldots, u_{1N}+n(\bar{x}_1-x_{10})(\bar{x}_N-x_{N0}) \\ \vdots \\ u_{N1}+n(\bar{x}_N-x_{N0})(\bar{x}_1-x_{10}), \ldots, u_{NN}+n(\bar{x}_N-x_{N0})^2 \end{vmatrix}$$

(H-40)

If $[u_{ij}]$ is a positive definite matrix, the above equation can be written as

$$N^{S}_{x_o,\,n} = |u_{ij}| \cdot \left[1 + n \sum_{i,j=1}^{N} u^{ij}(\bar{x}_i-x_{io})(\bar{x}_j-x_{jo})\right], \qquad \text{(H-41)}$$

where u^{ij} are elements of the inverse matrix $[u_{ij}]^{-1}$. The matrix $[u_{ij}]$ will be positive definite if and only if the n sample points do not lie in a flat space of less than N dimensions. From Equation H-41 it follows that the minimum value of the scatter $N^{S}_{x_o,\,n}$, occurs when the pivotal point $(x_{10}, x_{20}, \ldots, x_{N0})$ is chosen as the vector of sample means $(\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_N)$. This minimum value is then

$$N^{S}_{\bar{x},\,n} = |u_{ij}| \cdot \qquad \text{(H-42)}$$

The determinant $\left|u_{ij}\right|$ is called the internal scatter of the sample, and the matrix $\left[u_{ij}\right]$ is called the internal scatter matrix.

7. Linear Discriminant Analysis Based on Scatter-Two Groups

Let

$$x_{r_1}^{(1)} = (x_{1r_1}^{(1)}, \ldots, x_{Nr_1}^{(1)}),\ r_1 = 1,2, \ldots, n_1$$

and

$$x_{r_2}^{(2)} = (x_{1r_2}^{(2)}, \ldots, x_{Nr_2}^{(2)}),\ r_2 = 1,2, \ldots, n_2$$

be two samples of size n_1 and n_2, with $n_1 > N$ and $n_2 > N$. Let $x_i^{(1)}$ and $x_i^{(2)}$ be, respectively, the means of the x_i, $i = 1,2, \ldots, N$. for the two samples. Let $\left[u_{ij}^{(1)}\right]$ and $\left[u_{ij}^{(2)}\right]$ denote, respectively, the internal scatter matrices of the two samples with $u_{ij}^{(g)}$, $g = 1,2$, as defined in Equation H-35. Assume also that the internal scatter matrices are nonsingular. By pooling the two samples, one obtains a grand sample for which the sample means are denoted by $\bar{x}_i$ and the internal scatter matrix by $\left[u_{ij}\right]$. Now define the within-samples scatter matrix $W = \left[w_{ij}\right]$ by

$$W = \left[w_{ij}\right] = \left[u_{ij}^{(1)} + u_{ij}^{(2)}\right] \tag{H-43}$$

Note that if we assumed the covariance matrices of the two groups to be equal, the sample estimate of the common covariance matrix would be given by $\left[s_{ij}\right]$, where

$$(n_1+n_2-2)\, s_{ij} = \sum_{r=1}^{n_1} (x_{ir}^{(1)} - \bar{x}_i^{(1)})(x_{jr}^{(1)} - \bar{x}_j^{(1)}) + \sum_{r=1}^{n_2} (x_{ir}^{(2)} - \bar{x}_i^{(2)})(x_{jr}^{(2)} - \bar{x}_j^{(2)}) = w_{ij}\ . \tag{H-44}$$

Now the internal scatter matrix of the grand sample whose mean is x_i is $\left[u_{ij} \right]$, where

$$u_{ij} = u_{ij}^{(1)} + u_{ij}^{(2)} + n_1 (\bar{x}_i^{(1)} - \bar{x}_i)(\bar{x}_j^{(1)} - \bar{x}_j) + n_2 (\bar{x}_i^{(2)} - \bar{x}_i)(\bar{x}_j^{(2)} - \bar{x}_j) \tag{H-45}$$

so that the between-samples scatter matrix $B = \left[b_{ij} \right]$ has its elements given by

$$\left.\begin{aligned} b_{ij} &= u_{ij} - w_{ij} , \\ &= n_1 (\bar{x}_i^{(1)} - \bar{x}_i)(\bar{x}_j^{(1)} - \bar{x}_j) + n_2 (\bar{x}_i^{(2)} - \bar{x}_i)(\bar{x}_j^{(2)} - \bar{x}_j) \\ &= \frac{n_1 n_2}{n_1 + n_2} (\bar{x}_i^{(1)} - \bar{x}_i^{(2)})(\bar{x}_j^{(1)} - \bar{x}_j^{(2)}) \end{aligned}\right\} \tag{H-46}$$

Therefore, $b_{ij} = b_i b_j$, where $b_i = \sqrt{\frac{n_1 n_2}{n_1 + n_2}} (\bar{x}_i^{(1)} - \bar{x}_i^{(2)}) = \sqrt{\frac{n_1 n_2}{n_1 + n_2}}\, d_i$.

Let $A = (a_1, a_2, \ldots, a_N)$ be an arbitrary vector and let the two samples be denoted by $x^{(g)}$, $g = 1, 2$ so that

$$x^{(g)} = \begin{bmatrix} x_{11}^{(g)}, & x_{12}^{(g)}, & \ldots, & x_{1 n_g}^{(g)} \\ \vdots & \vdots & & \vdots \\ x_{N1}^{(g)}, & x_{N2}^{(g)}, & \ldots, & x_{N, n_g}^{(g)} \end{bmatrix} .$$

Projecting the two samples onto a line whose direction cosines are proportional to the elements of the vector A gives

$$Z'^{(g)} = A'^{(g)} x^{(g)}, \quad g = 1, 2 \tag{H-47}$$

and $Z^{(1)} = (z_1^{(1)}, \ldots, z_{n_1}^{(1)})$ and $Z^{(2)} = (z_1^{(2)}, \ldots, z_{n_2}^{(2)})$

are the two one-dimensional samples obtained as a result of the projection. Let $\bar{z}^{(1)}$ and $\bar{z}^{(2)}$ be the means of the two one-dimensional samples of z's and let $\bar{z}$ be the mean of the pooled sample. Then, if $_1S_{z,\, n_1+n_2}$ is the scatter of the grand sample obtained by pooling the two one-dimensional samples of z's, we can write

$$_1S_{\bar{z},\, n_1+n_2} = S_W + S_B \,, \tag{H-48}$$

where S_W is the within-samples scatter of the z's, given by

$$S_W = \sum_{r=1}^{n_1} (z_r^{(1)} - \bar{z}^{(1)})^2 + \sum_{r=1}^{n_2} (z_r^{(2)} - \bar{z}^{(2)})^2, \tag{H-49}$$

and S_B is the between-samples scatter of the z's, given by

$$\begin{aligned} S_B &= n_1(\bar{z}^{(1)} - \bar{z})^2 + n_2(\bar{z}^{(2)} - \bar{z})^2 \\ &= \frac{n_1 n_2}{n_1+n_2} (\bar{z}^{(1)} - \bar{z}^{(2)})^2 \end{aligned} \tag{H-50}$$

The approach of linear discriminant analysis based on scatter is to determine the particular value of $A = (a_1, a_2, \ldots, a_N)$ which maximizes S_B, the between-sample scatter for a fixed value of S_W, the within-sample scatter. Thus, the problem is to maximize S_B subject to the constraint S_W = constant.

Now,

$$\begin{aligned} S_B &= n_1(\bar{z}^{(1)} - \bar{z})^2 + n_2(\bar{z}^{(2)} - \bar{z})^2 \\ &= n_1\left[\sum_{i=1}^{N} a_i(\bar{x}_i^{(1)} - \bar{x}_i)\right]^2 + n_2\left[\sum_{i=1}^{N} a_i(\bar{x}_i^{(1)} - \bar{x}_i)\right]^2 \\ &= \sum_{j=1}^{N}\sum_{i=1}^{N} a_i a_j \left[n_1(\bar{x}_i^{(1)} - x_i)(\bar{x}_j^{(1)} - \bar{x}_j) + n_2(\bar{x}_i^{(2)} - \bar{x}_i)(\bar{x}_j^{(2)} - x_j)\right] \\ &= A'BA, \end{aligned} \tag{H-51}$$

where B is the matrix defined in H-46.

Similarly,

$$S_W = A'WA. \tag{H-52}$$

Using λ to denote a Lagrange multiplier, a necessary condition for a maximum is then

$$\frac{\partial}{\partial A}\left[S_B - \lambda \ (S_W - \text{const}) \right] = 0$$

or

$$2A'B - \lambda \ 2A'W = 0$$

or

$$A' \left[B - \lambda W \right] = 0. \tag{H-53}$$

For a non-zero value of A, the determinant of the term in brackets must be zero, i.e.,

$$\left| B - \lambda \ W \right| = 0 \tag{H-54}$$

or

$$\left| b_{ij} - \lambda \ w_{ij} \right| = 0$$

or

$$\left| b_i b_j - \lambda \ w_{ij} \right| = 0.$$

Then,

$$\lambda_1 = \sum \sum \ w^{ij} \ b_i b_j$$

$$= \frac{n_1 n_2}{n_1 + n_2} \ d' \ W^{-1} \ d, \tag{H-55}$$

where b_i is as defined in Equation H-46 and

$$d = \begin{bmatrix} d_1 \\ d_2 \\ d_3 \\ \cdot \\ \cdot \\ \cdot \\ d_n \end{bmatrix}$$

and

$$d_i = (\bar{x}_i^{(1)} - \bar{x}_i^{(2)}). \qquad \text{(H-56)}$$

The value of A corresponding to this root λ_1, satisfies the equation

$$\sum_{j=1}^{N} (b_{ij} - \lambda_1 W_{ij}) a_j = 0. \qquad \text{(H-57)}$$

The value of A can be determined within a constant of proportionality by noting that

$$S_B = \frac{n_1 n_2}{n_1 + n_2} (\bar{z}^{(1)} - \bar{z}^{(2)})^2 = \frac{n_1 n_2}{n_1 + n_2} \left[d'A \right]^2 ,$$

and

$$S_W = A' W A = (n_1 + n_2 - 2) A' S A,$$

where S is sample estimate of the common covariance matrix for the two groups. Then,

$$\frac{\partial}{\partial A} \left[S_B - \lambda (S - \text{const}) \right] = 0 \qquad \text{gives}$$

$$\frac{n_1 n_2}{n_1 + n_2} \cdot 2 (d'A) \cdot d' - \lambda \cdot 2 (n_1 + n_2 - 2) \cdot A'S = 0$$

or

$$\frac{n_1 n_2}{n_1 + n_2} \quad \frac{d'A}{\lambda (n_1 + n_2 - 2)} \cdot d' = A'S.$$

Since every element of d' is multiplied by the same term and since we can determine the elements of A only within a constant of proportionality, we set

$$\frac{n_1 n_2}{(n_1 + n_2) \lambda} \quad \frac{d'A}{(n_1 + n_2 - 2)} = \text{constant, say K.}$$

Then, noting that S is a symmetric matrix, from K.d' = A's,

$$A = K.\ S^{-1}\ d. \tag{H-58}$$

With K = 1 we get the same direction as with any other value of K. Using $A = S^{-1}\ d$, the linear function

$$x'\ A = x'\ S^{-1}\ d \tag{H-59}$$

which is obtained is Fishers' linear discriminant function. The boundary for classification is given by

$$x'\ S^{-1}\ d = \text{constant}. \tag{H-60}$$

Note that maximizing S_B while keeping S_W constant is the same as maximizing the ratio $\frac{S_B}{S_W}$ or minimizing the ratio $\frac{S_W}{S_B + S_W}$ while keeping S_W constant.

Alternative derivations of the above direction of projection have been given recently in the pattern recognition literature in terms of average Euclidean distance between pairs of points within groups, between groups, and among pooled samples. These "new" formulations are trivially different, do not lead to any new results, and are more cumbersome, since the average Euclidean distance between pairs of points does not have the simple relationship to spread of a sample as does the scatter about a pivotal point as formulated in multivariate statistical analysis and described above.

8. Multiple Linear Discriminant Functions

When there are G groups, G > 2, we can proceed to consider each pair of groups separately. However, it is possible to study overall relationships of the groups with a certain amount of parsimony by defining within-samples and between-samples scatter matrices for G Groups. The within-samples scatter matrix now has elements

$$w_{ij} = \sum_{g=1}^{G} \sum_{r=1}^{n_g} (x_{ir}^{(g)} - \bar{x}_i^{(g)})\,(x_{jr}^{(g)} - \bar{x}_j^{(g)}) \tag{H-61}$$

and the between-samples scatter matrix has elements

$$b_{ij} = \sum_{g=1}^{G} n_g\,(\bar{x}_i^{(g)} - \bar{x}_i)\,(\bar{x}_j^{(g)} - \bar{x}_j)\ . \tag{H-62}$$

The within-samples scatter matrix $[\, w_{ij} \,]$ is sometimes called the SSW, SPW matrix, the terms referring to "sums of squares within" and "sums of products within." Similarly, the matrix $[\, b_{ij} \,]$ is sometimes called the SSB, SPB matrix, these terms referring to "sums of squares between" and "sums of products between."

Proceeding as in the two-group case, one sets up the condition for maximizing the between-samples scatter of the projected samples while keeping the within-samples scatter constant. As in the two group case, this leads to the condition that

$$\left| W^{-1} \; B - \lambda \; I \right| \; = \; 0 \qquad \text{(H-63)}$$

Unlike the two-group case though, now there will be λ_j , $[j = 1, \ldots,$ min $(G-1, n)\,]$ roots which are non-zero, unless it so happens that all the sample group means lie along the same line. Corresponding to each root λ_j will be a vector A_j which is a solution of the equation

$$A'_j \; \left[\; B - \lambda_j \; W \; \right] \; = \; 0 \qquad \text{(H-64)}$$

Each A_j leads to a linear discriminant function $Z_j = \sum_{i=1}^{N} a_{ji} \, x_i$.

The samples may now be represented in this space, of reduced dimensions, which is referred to as the "discriminant space." Classification in the new space can then proceed according to procedures available. Since the new variables Z_j are weighted linear sums of the variables x_i, even when the x_i are discrete variables the approximation of multivariate normal distributions for the new variables is likely to prove useful.

9. Euclidean Distance as a Means for Classification of Projected Samples

Some investigators have proposed the use of the mean-squared Euclidean distance between pairs of points after projection of the samples (on a subspace obtained from the original N-dimensional space by a linear transformation) as a suitable measure by which to classify a new sample into one of two groups. It is easily shown that computing ordinary Euclidean distance in the new space does not result in a quadratic discriminant function but only in a linear function. Thus, the additional computation needed to obtain Euclidean distances can be avoided. (Only when the coordinates of the new space are suitably weighted, as, for example, when the functions defining the discriminant space are weighted

in a manner which accounts for their relative importance, does the computation of Euclidean distances prove useful.)

Let $Z = (z_1, z_2, \ldots, z_K)$ be the new sample after projection into the transformed space. Let $z_{r_1} = (z_{ir_1}^{(1)}, \ldots, z_{K,r_1}^{(1)})$, $r_1 = 1, 2, \ldots, n_1$ be the samples of Group 1 after projection and $z_{r_2}^{(2)} = (z_{1r}^{(2)}, \ldots, z_{K,r_2}^{(2)})$ $r_2 = 1, 2 \ldots n_2$ be the samples of Group 2 after projection. The average Euclidean distance between the new sample and points belonging to Group 1 is

$$\frac{1}{n_1} \sum_{r_1=1}^{n_1} \sum_{i=1}^{K} (z_i - z_{ir_1}^{(1)})^2 \qquad \text{(H-65)}$$

$$= \sum_{i=1}^{K} z_i^2 + \sum_{i=1}^{K} \overline{z_i^2}(1) - \sum_{i=1}^{K} \overline{z}_i(1) \cdot z_i \quad . \qquad \text{(H-66)}$$

Similarly, the average Euclidean between the new sample and points belonging to Group 2 is

$$\frac{1}{n_2} \sum_{r_2=1}^{n_2} \sum_{i=1}^{K} (z_i - z_{ir_2}^{(2)})^2$$

$$= \sum_{i=1}^{K} z_i^2 + \sum_{i=1}^{K} \overline{z_i^2}(2) - \sum_{i=1}^{K} \overline{z}_i(2) \cdot z_i \qquad \text{(H-67)}$$

Taking the difference between the two gives the expression

$$\sum_{i=1}^{K} (\overline{z}_i^{(2)} - \overline{z}_i^{(1)}) \overline{z}_i + \sum_{i=1}^{K} (\overline{z_i^2}(1) + \overline{z_i^2}(2)) \qquad \text{(H-68)}$$

as the classification procedure. This expression is a linear function in the z_i. Furthermore, the coefficients of the z_i are simply the difference between the mean values of the projected coordinates in the two groups, the projected coordinates having been obtained by a linear transformation from the variables x_i. Thus no quadratic discriminant function results.

10. Mahalanobis' D^2 *, **

For two samples of size n_1 and n_2 with N characteristics measured on each member of each sample

$$D_N^2 = \sum_{i=1}^{N} \sum_{j=1}^{N} s^{ji} (\bar{x}_i^{(1)} - \bar{x}_i^{(2)}) (\bar{x}_j^{(1)} - \bar{x}_j^{(2)}), \qquad \text{(H-69)}$$

where the matrix $[s^{ji}]$ is the inverse of $[s_{ij}]$,

$$x_i^{(g)} = \frac{1}{n_g} \sum_{r=1}^{n_g} x_{ir}^{(g)}, \quad i = 1, 2, \ldots, N; \; g = 1, 2$$

and

$$(n_1 + n_2 - 2)\, s_{ij} = \sum_{g=1}^{2} \sum_{r=1}^{n_g} (x_{ir}^{(g)} - \bar{x}_i^{(g)}) (x_{jr}^{(g)} - \bar{x}_j^{(g)}) \qquad \text{(H-70)}$$

$$= w_{ij},$$

where $[w_{ij}]$ is the within-samples scatter matrix defined in Equation H-44. In terms of

$$d = \begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ d_N \end{bmatrix} \qquad \text{with } d_i = (\bar{x}_i^{(1)} - \bar{x}_i^{(2)}),$$

$$D_N^2 = (n_1 + n_2 - 2)\, d' W^{-1} d. \qquad \text{(H-71)}$$

Hottelling's T^2, a generalization of Student's t for two multivariate samples, is related to D^2 by

$$T^2 = \frac{n_1 n_2}{n_1 + n_2} D_N^2 . \qquad \text{(H-72)}$$

---

* Rao, C.R., Advanced Statistical Methods in Biometric Research, New York, John Wiley, 1952, Chap. 7.

** Wilks, S.S., Mathematical Statistics, New York, John Wiley and Sons, 1962, Chap. 18.

To test the hypothesis that there is no difference in the mean values of the N characteristics for the two populations, the statistic

$$\frac{(n_1+n_2-N-1)}{(n_1+n_2-2)\,N}\cdot T^2 = \frac{n_1 n_2}{(n_1+n_2)}\;\frac{(n_1+n_2-N-1)}{(n_1+n_2-2)\,N}\cdot D_N^2 \qquad \text{(H-73)}$$

is used. For the case where both samples are independently drawn from identical N-dimensional normal distributions, i.e., under the null hypothesis, the statistic of Equation H-73 has an F distribution with N and $(n_1 + n_2 - N - 1)$ degrees of freedom.

For G samples, G > 2, of size n_g, g = 1, 2, ..., G, Rao's generalization of Mahalanobis' D^2 is given by

$$D^2_{N,G} = \sum_{i=1}^{N}\sum_{j=1}^{N} s^{ij} \sum_{g=1}^{G} n_g \left(\bar{x}_i^{(g)} - \bar{x}_i\right)\left(\bar{x}_j^{(g)} - \bar{x}_j\right), \qquad \text{(H-74)}$$

where

$$\left(\sum_{g=1}^{G} n_g - 1 - G\cdot N\right) s_{ij} = \sum_{g=1}^{G}\sum_{r=1}^{n_g} \left(x_{ir}^{(g)} - \bar{x}_i^{(g)}\right)\left(x_{jr}^{(g)} - \bar{x}_j^{(g)}\right)$$

$$= w_{ij}$$

and

$$x_i = \frac{\sum_{g=1}^{G} n_g \bar{x}_i^{(g)}}{\sum_{g=1}^{G} n_g}$$

Equation H-74 can be rewritten as

$$D^2_{N,G} = \left(\sum_{g=1}^{G} n_g - 1 - G\cdot N\right) \sum_{j=1}^{N}\sum_{i=1}^{N} w^{ji}\, b_{ij}$$

$$= \left(\sum_{g=1}^{G} n_g - 1 - G\cdot N\right) \text{trace } W^{-1} B. \qquad \text{(H-75)}$$

When the total sample size $\sum_{g=1}^{G} n_g$ is large, then $D^2_{N,G}$ as defined in Equation H-75 has a χ^2 distribution with N(G-1) degrees of freedom under the null hypothesis of no difference between the mean values of the N characteristics in all the G populations. This statistic can be used for testing whether observations on Q additional variables can be useful for increasing the distance between the two samples. Rao suggests the use of $(D^2_{N+Q,G} - D^2_{N,G})$ to judge the significance of information supplied by the additional Q variables. For very large sample sizes this difference has, approximately, a χ^2 distribution with Q(G-1) degrees of freedom. As in the case of two groups, a step-by-step screening of the variables can be done using this statistic.

11. The Symmetric Divergence for Binary Variables

A general measure of the divergence between two populations is obtained by taking the difference between expected values of log L(x) with respect to the probability functions of Group 1 and Group 2, respectively. Thus,

$$J(1,2) = E_{f_1}\left[\log \frac{f_1(x)}{f_2(x)}\right] - E_{f_2}\left[\log \frac{f_1(x)}{f_2(x)}\right]. \qquad \text{(H-76)}$$

J is also called the Kullback-Leibler information number. Note that for the case where the two populations are multivariate normal with different mean vectors but equal covariance matrices, the symmetric divergence, J, gives Mahalanobis' D^2 for two populations, i.e.,

$$D^2 = \sum_{i=1}^{N}\sum_{j=1}^{N} v^{ji}\,(m_i^{(1)} - m_i^{(2)})\,(m_j^{(1)} - m_j^{(2)}). \qquad \text{(H-77)}$$

For the case where the variables $x_1, x_2, \ldots, x_N$ are binary, in terms of the notation of Section 4 of this Appendix, an approximation to the effective distance between the two groups is given by*

* Bahadur, R.R., "On Classification Based on Responses to N Dichotomous Items," USAF SAM Series in Statistics, Randolf AFB Texas 1959; appears in Studies in Item Analysis and Prediction, Ed. Herbert Solomon, Stanford U. Press, 1961.

$$J \approx D^{*2} = \sum_{i=1}^{N} (m_i^{(1)} - m_i^{(2)})^2 a_i + \sum_{i<j}^{N} (r_{ij}^{(1)} - r_{ij}^{(2)})^2$$

$$+ \ldots + (r_{12 \ldots N}^{(1)} - r_{12 \ldots N}^{(2)})^2 , \qquad \text{(H-78)}$$

where a_i is as defined in Equation H-22. D^* does not satisfy the triangle inequality and so is not a metric. It represents a "good" approximation to the effective distance when the two population distributions are not very different and the predictors are not highly correlated within either group.* A rough idea of the usefulness of a set of predictors for discrimination can be obtained by using the first two terms for D^{*2}.

---

* Bahadur, R. R., "On Classification Based on Responses to N Dichotomous Items," USAF SAM Series in Statistics, Randolph AFB, Texas 1959; appears in Studies in Item Analysis and Prediction, Ed. Herbert Solomon, Stanford U. Press, 1961.

APPENDIX I

BIOGRAPHIES

LAVEEN N. KANAL, RESEARCH MANAGER

Dr. Kanal received the B.S. degree in electrical engineering in 1951 and the M.S. degree in electrical engineering in 1953, both from the University of Washington. His master's thesis is entitled Resonance Ionization and the Detection of Mercury Vapor. He studied under a Canadian General Electric Company advanced technical course from 1953 until 1954 and studied pre-doctoral courses at the University of Toronto from 1954 until 1955. He received the Ph.D. degree in electrical engineering from the University of Pennsylvania in 1960. The title of his doctoral thesis is Analysis of Some Stochastic Processes Arising from a Learning Model. The University of Washington awarded him a foreign exchange scholarship from 1949 to 1951 and a teaching fellowship from 1951 until 1953. He studied under a National Science Foundation grant from 1959 until 1960.

Dr. Kanal's specialties are pattern recognition, stochastic learning models, and applications of statistical theory to communication and control.

Dr. Kanal joined the Philco Scientific Laboratory in 1962 and is Research Manager of the Advanced Pattern Recognition Group. His work has included contractual studies of Semiautomatic Imagery Screening for Object Recognition and of Advanced Pattern Recognition. He has participated in writing proposals on advanced recognition techniques.

Dr. Kanal is also currently a special lecturer in Operations Research at the Wharton School of the University of Pennsylvania.

From 1960 to 1962, he was the Manager of the Pattern Recognition and Learning Laboratory of the Communications and Electronics Advanced Development Department of General Dynamics/Electronics. There he worked on machine recognition of patterns and speech using various statistical distance functions, an adaptive filter for signal detection and extraction, behavioral models for signal-detection by human observers, and reinforcement models for learning. Under his supervision a special purpose computer for experiments in self-organizing networks was designed and a programmer simulated some specific nets and demonstrated success in recognizing multi-font characters, isolated words, and in classifying vowels.

From 1955 to 1960, Dr. Kanal was a research engineer, instructor and associate on the staff at the Moore School of Electrical Engineering of the University of Pennsylvania. He participated in sponsored research projects on an electronic dial telephone system, contributed studies in artificial intelligence, taught undergraduate courses in network analysis and synthesis and electron tubes, and courses in probability theory, theory of control,

servomechanisms, and tube and transistor networks to graduate students in electrical engineering, systems engincering, and operations research. He also was associated with the Psychology Department.

From 1954 until 1955, Dr. Kanal was a Development Engineer at the Canadian General Electric Company. There he developed pulse circuits, filters, and wideband i-f amplifiers. From 1953-54 he participated in the C.G.E. "Test" program with assignments in the departments of motors and generators, circuit breakers, radio and television receivers, and electronic laboratory instruments.

From 1951 until 1953, Dr. Kanal was a Teaching Fellow at the University of Washington where he taught undergraduate electronics and power laboratory courses.

Dr. Kanal's publications include:

"Detection of Mercury Vapor by Resonance Ionization," presented at the summer meeting of the American Physical Society, June 1954. Abstract appears in the Physical Review for 1954.

"On Miyata's Method of Synthesis," letter to the editor, IRE Transactions on Circuit Theory, Vol. CT-4, No. 4, December 1957, p. 340-341.

"A Functional Equation Analysis of a Nonlinear Stochastic Model for Learning," presented at the Operations Research Society of America, Detroit meeting, October 1960.

"A Random Walk Related to a Nonlinear Learning Model," IRE Convention Record, Vol. 9, Part 2, March 1961, p. 211-216.

"A Functional Equation Analysis of Two Learning Models," Psychometrika, Vol. 27, No. 1, March 1962, pp. 89 to 104 (also included in Readings in Mathematical Psychology, John Wiley and Sons, 1963).

"Asymptotic Distribution for the Two-Absorbing - Barrier Beta Learning Model," Psychometrika, Vol. 27, No. 1, March 1962, pp. 105 to 109 (also included in Readings in Mathematical Psychology, John Wiley and Sons, 1963).

"Evaluation of a Class of Pattern Recognition Networks," presented at the second annual Bionics Symposium, Ithaca, New York, August 1961; appears in Biological Prototypes and Synthetic Systems, Plenum Press, 1962.

"An Analysis of a Class of Pattern Recognition Networks," presented at the Los Angeles meeting of the Instrument Society of America, September 1961.

"Basic Principles of Some Pattern Recognition Systems," Proceedings of the National Electronics Conference, Chicago 1962 (with others).

Dr. Kanal has two patents pending, both for a pattern recognition system.

Dr. Kanal is a member of the Discrete Systems Theory Committee of the IEEE, being Co-Chairman of the 1st Symposium on Discrete Adaptive Processes at the 1962 Joint Automatic Control Conference and Chairman of the 2nd Symposium on Discrete Adaptive Processes in 1963. He is also a member of Sigma Xi, the Institute of Mathematical Statistics, and the American Statistical Association. He is a Registered Professional Engineer in the Province of Ontario, Canada.

WOLFGANG F. WERNER, RESEARCH SPECIALIST

Dr. Werner received the M.S. degree in theoretical physics in 1958 and the Ph.D. degree in applied mathematics in 1960, both from the Technical University of Munich. His specialties in graduate school included partial differential equations, numerical analysis, and theoretical gas dynamics. His master's thesis is entitled A Limiting Case in the Theory of X-Ray Impulse Radiation, and his doctoral dissertation, Instability of Transonic, Shockless Flow Past Profiles.

Dr. Werner's principal fields of interest are numerical analysis and theoretical mechanics, with emphasis on numerical solutions of three-dimensional partial differential equations.

Dr. Werner joined the Philco Scientific Laboratory in 1961 to serve as a consultant in mathematics, mathematical physics, and numerical analysis, and to aid in the research of the logics of pattern recognizers.

Dr. Werner served as a member of the research staff of the Mathematics Institute of the Technical University of Munich for three and one half years. He also lectured in mathematics at the University of Maryland, European Extension.

Dr. Werner is the author of the following papers:

"A Method for Calculation of Flow Fields Behind the Bow Shock of Plane or Rotational Symmetric Bodies," to be published in Zeitschrift fur Angewandte Mathematik und Mechanik.

"Solution of a Problem in the Theory of Thin Plates and its Numerical Computation," to be published in Zeitschrift fur Angewandte Mathematik und Mechanik.

"Analysis of Error Rates in Pattern Recognition," Journal of the Association of Computing Machinery, August 1962.

"Instability of Continuous Transonic Flow Past Profiles," Zeitschrift fur Angewandte Mathematik und Mechanik, Vol. 41, November 1961, pp. 448 to 458.

"Calculations of Shockfree Transonic Flows and Investigations on Stability of Such Solutions," Deutsche Forschingsgemeinshaft, (Federal Agency promoting scientific research) DFG Sa/19/28, January 1960.

A Numerical Approach to the Question of Stability of Enclosures in an Otherwise Subsonic Flow, Air Force Office of Scientific Research, Contract No. AF 61(514)-1080, November 1959.

"Numerical Calculations of Transonic Flow around Plane Profiles," Deutsche Forschingsgemeinshaft (Federal Agency promoting scientific research) DFG 12/19/27, July 1959.

On the Validity 'Area Rule' and the Theorem of Equivalence in Supersonic Flow, Air Force Office of Scientific Research, Contract No. AF 61(514)-1080, February 1958.

DONALD R. TAYLOR, JR., RESEARCH SPECIALIST

Mr. Taylor received the B.S. degree in electrical engineering from the Drexel Institute of Technology in 1950. In addition, Mr. Taylor studied under an eleven-month radar training course in the U.S. Navy, and has taken graduate level courses at the University of Pennsylvania and UCLA. As an undergraduate, he worked in factory quality control, radar engineering and development, and in radar research.

Mr. Taylor joined Philco on a full-time basis in June 1950. He has designed and developed linear and logarithmic i-f amplifier circuits and low-noise input stages for low-noise radar receivers and has published on this subject. He has flight tested experimental radar systems and worked on the missile guidance electronics.

From March 1954 until November 1957, Mr. Taylor was engaged in circuit research to avoid patent infringement in commercial radio and television products; and developed circuits later used in production. He pioneered in developing transistor circuits for television, for which he holds several patents and was responsible for research design of receiver circuits in Philco's first laboratory all-transistor set. This involved transistorization of video amplifiers circuits; deflection oscillators, drivers, and outputs; audio stages, sound i-f amplifiers and limiters; sync separators and AFC synchronizing circuits; high-voltage generation; and systems design. The advanced circuit research culminated in an eight-inch portable set in 1957 and provides the groundwork for the Philco "Safari" portable produced in 1959.

He has invented and tested novel circuit techniques in which multigrid tubes were used to combine receiver functions, and supervised experiments in which tunnel diodes were used as uhf oscillators. Several patent applications were made on the basis of this work.

As a result of a 2 year stay in Norway, Mr. Taylor also has experience in the design, development and production of European-Standard television receivers.

Since 1961, Mr. Taylor has been engaged in research on Information Storage and Retrieval and Artifical Intelligence. His recent activities have been in implementation studies for visual image processing devices, pattern recognition systems, and adaptive majority logic. He holds two patents. Eight patent applications filed since 1960 are pending issue. They are principally concerned with television techniques and novel circuit and device inventions for signal handling and pulse processing systems.

He is a member of Tau Beta Pi, Eta Kappa Nu, Phi Kappa Phi, and the American Documentation Institute.

JOHN Z. GRAYUM, RESEARCH SPECIALIST

Mr. Grayum received the B.S. degree in physics from St. Joseph's College in 1951 and the M.A. degree in physics from Temple University in 1958. He has done graduate work in electrical engineering, mathematics, and physics at Temple University and the Drexel Institute of Technology. He now is doing graduate work in mathematics at the University of Pennsylvania.

After he joined Philco in 1955, Mr. Grayum participated in the development of a transistorized rc active audio filter. He also helped develop passive linear time-delay filters and delay-line filters.

Mr. Grayum has directed work on REentrant DAta Processors (REDAP) and has made detailed system analyses of sweep integrators and iterators. Later, he directed a group of Industrial and Computer Laboratory engineers in work on PCM multiplex equipment, high-speed switching circuits, satellite transmitters, and digital data-handling equipment.

In 1958 he began conducting and directing studies of model communications, especially as related to secure communications systems. This work provided significant improvements in encoding, detection and decision and sync-scanning procedures.

Mr. Grayum was project leader of studies in the design of a world wide communication system using large passive satellites as reflectors. These studies included a consideration of A/J technicues. He was the project leader, and principle contributor of advanced studies to determine methods of reducing the vulnerability of active satellite communications links to jamming and spoofing by an intelligent opponent. On a Philco sponsored program, he directed studies of A/J techniques including "Methods of Spreading High Speed Data," "Coding of Higher Order Alphabets," "Variable Mode Communications," "The Evaluation and Comparison of F-T Dodging with the Philco Hybrid Techniques" and "Methods of Sync Search."

More recently Mr. Grayum has participated in fundamental studies of time-variable communication systems, including variable data rate and feedback techniques, and sequential decision procedures, including hypothesis testing and ranking.

From 1951 to 1955, Mr. Grayum was a Broadcast Engineer at RCA, where he contributed to the design and development of UHF broadcast antennas and filters.

Mr. Grayum is the author of "Optimum Decision and Scanning Techniques for Synchronization," Natcom Symposium Record, October 1-3, 1962, pp. 170-178.

He is the coauthor of an oral presentation: "Advanced Spread Spectrum Studies," presented at the National Security Agency, Fort Meade, Maryland, January 1961; also presented at CCDD, Bedford, Mass., May 1961, and at the Air Force Cambridge Research Laboratories, Lexington, Mass., June 1961, (coauthor: C. Gumacos).

Mr. Grayum is a member of Sigma Pi Sigma, a Senior Member of the IEEE and is listed in Who's Who in American Universities and Colleges.

CONSTANTINE GUMACOS, RESEARCH SPECIALIST

Mr. Gumacos received the B.S.E.E. degree with highest honors from the Georgia Institute of Technology in 1951 and has done graduate work at the Moore School of Electrical Engineering at the University of Pennsylvania.

After he joined Philco in 1951, Mr. Gumacos worked on sweep integrators for the SG-6/SG-6b radar equipment and on microwave relay systems for microwave and color-television communications. He also worked on low-noise cooled crystal receivers. In other projects, he was responsible for the design of a data processor for a coherent airborne radar system with high azimuthal resolution and worked on the analysis of a velocity-shaped MTI data-processing system.

From October 1958 to January 1960, Mr. Gumacos contributed to advanced studies of the Spread Eagle System, a jam-resistant, data link communications system, after which he participated in the systems design for a program for a global communication network which uses passive spherical satellites. Later, he participated in the "Midas Command Link Modulation" study and the "Midas Command Link Reliability" study.

Mr. Gumacos has done extensive work in the fields of coding and synchronization for A/J communication systems. He participated in studies of codes for secure communications, using the concepts of modern algebra. Most recently he performed a study of the theoretical aspects of synchronizing secure systems. He has also performed numerous analyses of modulation techniques.

Mr. Gumacos has published these papers:

"Advanced Coding Studies," Internal Philco Report, June 1960.

"Analysis of an Optimum Sync Search Procedure," submitted to the IRE Transactions on Communication Systems.

"Analysis of Multiple Frequency Shift Keying with Diversity," Philco Internal Report, September 15, 1961 (coauthor: Peter M. Hahn).

"Advanced Spread Spectrum Studies," presented at the National Security Agency, Fort Meade, Maryland, January 1961; also presented at CCDD, Bedford, Mass., May 1961, and at the Air Force Cambridge Research Laboratories, Lexington, Mass., June 1961 (coauthor: J. Z. Grayum).

HENRY G. KELLETT, SENIOR ENGINEER

Mr. Kellett is a member of the Image Recognition Group of the Advanced Technology Laboratory where he is engaged in the study of problems in implementing an automatic imagery screening system. He is also engaged in the development of linear spatial filtering techniques for prenormalization and screening of aerial reconnaissance data. After joining Philco in 1959 he worked on the Data Conversion Program in which he was responsible for the design and development of the encoder portion of the equipment. For several months, he also contributed to research on secure, jam resistant, and private communication systems, and has recently been involved in advanced object recognition studies.

Mr. Kellett received the B.S. degree in electrical engineering from the University of New Hampshire in 1959, and now is doing graduate work in electrical engineering at the University of Pennsylvania. In addition to his formal education, Mr. Kellett has attended courses in electronics in the Air Force and at the evening school of the Georgia Institute of Technology.

Prior to obtaining his Baccalaureate degree, Mr. Kellett was employed in several technical positions in industry and in the Air Force.

He is a member of the IEEE and its Professional Group on Information Theory.

JERRY R. RICHARDS, SENIOR ENGINEER

Mr. Richards is a member of the Image Recognition Group of the Advanced Technology Laboratory where he is working on studies and computer simulations of the logics used in pattern-recognition systems. Recently, Mr. Richards investigated a logic for the detection of significant changes in information in repeat-cover aerial photography. His work in this area includes the preparation of computer simulation programs for the recently concluded Automatic Video Data Analysis and Image Change Detection projects. Mr. Richards also has served as a consultant on computer simulation of learning-machine logic on a project investigating adaptive pattern-recognition techniques.

Mr. Richards was granted the B.S. degree in electrical engineering by the Drexel Institute of Technology in June 1959.

During his student years, Mr. Richards worked as a Cooperative Student Engineer on such projects as the design of electrical furnaces and the optimization of electrolysis. Since graduation, he has been a member of the Philco Corporation. He has helped to design gamma-correcting circuits, has evaluated photographic material for use in a project to reduce the redundancy of pictorial data before transmission, has studied the characteristics of switching thin-magnetic films, and has participated in the design of the video processor for an electronic, variable-font address reader being built for the U.S. Post Office Department. He also has studied video-processing techniques as applied to aerial photographs.

Mr. Richards is a Member of the IEEE.

JEROME I. MANTELL, PHYSICIST

Mr. Mantell received the B.A. degree in physics from the University of Pennsylvania in 1961. He now is working toward the master's degree at the same university.

In August 1960, Mr. Mantell joined the Philco Research Division, after which he performed experimental and theoretical studies on the field properties of magnesium oxide, with emphasis on the mechanism of secondary electron emission. The studies included investigation of the properties of magnesium oxide in air and vacuum. Mr. Mantell prepared several patent disclosures for a magnesium oxide, single crystal vacuum tube employing cold-cathode emission. He constructed a simple theoretical model of the secondary emission phenomenon in terms of field effects.

From November 1960 to February 1961 Mr. Mantell did experimental work in which an electrochemical light valve was used as a display and as an integral part of a logic network including studies of various properties of the electroplating cell, including such properties as efficiency, transmission of light, etc. He also conducted studies into preparation of a bistable mode of operation.

Since February 1961, Mr. Mantell has performed theoretical experiments for optical logic systems, including programming, design and construction. Mr. Mantell now is investigating and constructing electroluminesent-photocell distable circuits and is working on a preliminary investigation of new kinds of circuit elements for learning machines.

Mr. Mantell is a Member of the American Physical Society and Pi Sigma.

HANS P. DOMABYL, ENGINEER

Hans P. Domabyl received the "Diplomingenier" degree in Electrical Engineering from the Munich Institute of Technology in 1957 and was granted a similar degree in Economics in 1959 after postgraduate studies in Paris and Munich.

He worked for two years as an application and sales engineer for the Ampex Corporation in Germany, where he was concerned with problems of adapting magnetic tape recorders for special instrumentation and telemetry systems. Mr. Domabyl joined the Philco Scientific Laboratory in 1962 and received training in programming the Philco 2000 computer system. At the present time, he is writing computer programs for automatic data recognition.

Mr. Domabyl is a member of the VDE, the German Institute of Electrical Engineers.

Thomas J. B. Shanley, Manager, Recognition Laboratory

Dr. Shanley received the B.S. degree in engineering from the U. S. Military Academy in 1939. He did graduate work in cosmic ray physics at Princeton University, which granted him the Ph.D. degree in 1951. The title of his doctoral thesis is, Gamma Ray Production in Mu Mesa Capture.

Dr. Shanley joined the Philco Corporation in 1961 and is now directing the research and advanced development work in the Advanced Technology Laboratory on visual pattern recognition and speech recognition.

From 1939 until he joined Philco, Dr. Shanley served in the U.S. Army. He rose from Second Lieutenant to Colonel during his Army career. He commanded a company in the original Army parachute unit, the 501st Parachute Battalion, from 1939 to 1946. During this period, he pioneered in the development of mass airborne drop techniques. In 1943, he was promoted to Lieutenant Colonel to command a parachute infantry battalion of the 82nd Airborne Division in the airborne assault of Normandy and commanded a regiment in combat. He received a Silver Star, two Bronze Stars, and the Purple Heart.

From 1950 to 1953, Dr. Shanley was the Chief of the Research Division in the Research and Development Section of the Office of the Chief of Army Field Forces. In this capacity, he formulated requirements for new Army weapons systems and conducted and reviewed studies of the effectiveness of weapons. He originated the "Combat Development" concept adopted by the Army in 1953, by which weapons, doctrine, and organizational concepts are conceived, developed, and tested concurrently.

From 1954 to 1955, Dr. Shanley commanded the 19th Infantry Regiment, 24th Infantry Division, in Korea. From 1958 to 1960, he was the Chief of the Atomic-Chemical and Biological Warfare Division in the Office of the Deputy Chief of Staff for Military Operations, Department of the Army Headquarters. In this capacity, he directed a group of Army officers in determining requirements for and planning for the use of nuclear delivery systems and chemical and biological weapons. In 1960, Dr. Shanley was assigned to the Office of the Joint Chiefs of Staff. He established requirements and prepared plans for the use of nuclear weapons and guided missiles.

Dr. Shanley's publications include the following:

"Influence on the Cosmic Ray Spectrum of Five Heavenly Bodies," Review of Modern Physics, Vol. 21, No. 1, January 1949, pp. 51 to 71 (coauthors: J. A. Wheeler and E. O. Kane).

"A Preliminary Directional Study of Cosmic Rays at High Altitude, II, Physical Review, Vol. 76, No. 8, October 15, 1949, pp. 1005 to 1019 (coauthors: J. R. Winckler and W. G. Stroud).

"Gamma Rays from Negative Mu Meson Capture in Lead," Physical Review, Vol. 89, No. 5, March 1953, pp. 983 to 990 (coauthor: G. G. Harris).

"Evaluation of Weapons, Tactics, and Organizational Concepts," Military Review, July 1954, pp. 31 to 36.

"Non-Nuclear NATO Army," Army, Vol. 11, No. 5, December 1960, pp. 28 to 30.

Dr. Shanley is a member of Sigma Xi, the Operations Research Society, the Association of the U. S. Army, and the IEEE.

Philco Advanced Technology Laboratories, Blue Bell, Pa., SEMI-AUTOMATIC IMAGERY SCREENING STUDY AND EXPERIMENTAL INVESTIGATION, by T. J. Harley, et. al., Second and Third Combined Quarterly Progress Report, 1 September 1962 to 28 February 1963, 216 pp., incl. illus., 62 refs. (Contract No. DA-36-039-SC-90742) Philco No. V043-2 and 3, UNCLASSIFIED REPORT

This report covers the work performed under Contract DA-36-039-SC-90742, "Semi-Automatic Imagery Screening," during the six-month period 1 September 1962 through 28 February 1963, comprising the fourth through ninth months of a 21-month program.

The following subjects are presented herein:

(a) Discussion of the merits of available techniques that may be considered for use in each stage of the process of screening aerial reconnaissance photographs for purposes of detecting objects of military significance.

(b) Description and comparison of alternative hardware realizations of the techniques of (a) above.

(c) Summary of the application of the principles of statistical decision theory to the reconnaissance photoscreening problem.

(d) Description of the conceptual design of an imagery screening device.

(e) Detailed description of a computer simulation problem on the detection of tanks in aerial photographs. The simulation, to be carried out during the next six months, is expected to provide design details to *extend the conceptual design of item (d).*

UNCLASSIFIED

1. Data Processing Systems
2. Photographic Interpretation
3. Target Recognition
I. Title: Imagery Screening
II. Harley, T. J., et.al.
III. United States Army Electronic Research and Development Laboratory, Fort Monmouth, New Jersey.
IV. Contract DA-36-039-SC-90742

UNCLASSIFIED

Philco Advanced Technology Laboratories, Blue Bell, Pa., SEMI-AUTOMATIC IMAGERY SCREENING STUDY AND EXPERIMENTAL INVESTIGATION, by T. J. Harley, et. al., Second and Third Combined Quarterly Progress Report, 1 September 1962 to 28 February 1963, 216 pp., incl. illus., 62 refs. (Contract No. DA-36-039-SC-90742) Philco No. V043-2 and 3, UNCLASSIFIED REPORT

This report covers the work performed under Contract DA-36-039-SC-90742, "Semi-Automatic Imagery Screening," during the six-month period 1 September 1962 through 28 February 1963, comprising the fourth through ninth months of a 21-month program.

The following subjects are presented herein:

(a) Discussion of the merits of available techniques that may be considered for use in each stage of the process of screening aerial reconnaissance photographs for purposes of detecting objects of military significance.

(b) Description and comparison of alternative hardware realizations of the techniques of (a) above.

(c) Summary of the application of the principles of statistical decision theory to the reconnaissance photoscreening problem.

(d) Description of the conceptual design of an imagery screening device.

(e) Detailed description of a computer simulation problem on the detection of tanks in aerial photographs. The simulation, to be carried out during the next six months, is expected to provide design details to extend the conceptual design of item (d).

UNCLASSIFIED

1. Data Processing Systems
2. Photographic Interpretation
3. Target Recognition
I. Title: Imagery Screening
II. Harley, T. J., et.al.
III. United States Army Electronic Research and Development Laboratory, Fort Monmouth, New Jersey.
IV. Contract DA-36-039-SC-90742

UNCLASSIFIED

Philco Advanced Technology Laboratories, Blue Bell, Pa., SEMI-AUTOMATIC IMAGERY SCREENING STUDY AND EXPERIMENTAL INVESTIGATION, by T. J. Harley, et. al., Second and Third Combined Quarterly Progress Report, 1 September 1962 to 28 February 1963, 216 pp., incl. illus., 62 refs. (Contract No. DA-36-039-SC-90742) Philco No. V043-2 and 3, UNCLASSIFIED REPORT

This report covers the work performed under Contract DA-36-039-SC-90742, "Semi-Automatic Imagery Screening," during the six-month period 1 September 1962 through 28 February 1963, comprising the fourth through ninth months of a 21-month program

The following subjects are presented herein:

(a) Discussion of the merits of available techniques that may be considered for use in each stage of the process of screening aerial reconnaissance photographs for purposes of detecting objects of military significance.

(b) Description and comparison of alternative hardware realizations of the techniques of (a) above.

(c) Summary of the application of the principles of statistical decision theory to the reconnaissance photoscreening problem.

(d) Description of the conceptual design of an imagery screening device.

(e) Detailed description of a computer simulation problem on the detection of tanks in aerial photographs. The simulation, to be carried out during the next six months, is expected to provide design details to extend the conceptual design of item (d).

UNCLASSIFIED

1. Data Processing Systems
2. Photographic Interpretation
3. Target Recognition
I. Title: Imagery Screening
II. Harley, T. J., et.al.
III. United States Army Electronic Research and Development Laboratory, Fort Monmouth, New Jersey.
IV. Contract DA-36-039-SC-90742

UNCLASSIFIED

Philco Advanced Technology Laboratories, Blue Bell, Pa., SEMI-AUTOMATIC IMAGERY SCREENING STUDY AND EXPERIMENTAL INVESTIGATION, by T. J. Harley, et. al., Second and Third Combined Quarterly Progress Report, 1 September 1962 to 28 February 1963, 216 pp., incl. illus., 62 refs. (Contract No. DA-36-039-SC-90742) Philco No. V043-2 and 3, UNCLASSIFIED REPORT

This report covers the work performed under Contract DA-36-039-SC-90742, "Semi-Automatic Imagery Screening," during the six-month period 1 September 1962 through 28 February 1963, comprising the fourth through ninth months of a 21-month program.

The following subjects are presented herein:

(a) Discussion of the merits of available techniques that may be considered for use in each stage of the process of screening aerial reconnaissance photographs for purposes of detecting objects of military significance.

(b) Description and comparison of alternative hardware realizations of the techniques of (a) above.

(c) Summary of the application of the principles of statistical decision theory to the reconnaissance photoscreening problem.

(d) Description of the conceptual design of an imagery screening device.

(e) Detailed description of a computer simulation problem on the detection of tanks in aerial photographs. The simulation, to be carried out during the next six months, is expected to provide design details to extend the conceptual design of item (d).

UNCLASSIFIED

1. Data Processing Systems
2. Photographic Interpretation
3. Target Recognition
I. Title: Imagery Screening
II. Harley, T. J., et.al.
III. United States Army Electronic Research and Development Laboratory, Fort Monmouth, New Jersey.
IV. Contract DA-36-039-SC-90742

UNCLASSIFIED

DISTRIBUTION LIST

Contract No. DA 36-039-SC-90742 Philco Corporation

Second and Third Combined Quarterly Report

Address	No. of Copies
OASD (R and E), Rm 3B1065 ATTN: Technical Library The Pentagon Washington 25, D. C.	1
Chief of Research and Development OCS, Department of the Army Washington 25, D. C.	1
Commanding General U. S. Army Materiel Command ATTN: R and D Directorate Washington 25, D. C.	1
Commanding General U. S. Army Electronics Command ATTN: AMSEL-CB Fort Monmouth, N. J.	2
Commander, Armed Services Technical Information Agency ATTN: TIPCR Arlington Hall Station Arlington 12, Virginia	10
Commanding General U. S. Army Electronics Research and Development Activity ATTN: Technical Library Fort Huachuca, Arizona	1
Chief, U. S. Army Security Agency Arlington Hall Station Arlington 12, Virginia	1
Deputy President U. S. Army Security Agency Board Arlington Hall Station Arlington 12, Virginia	1

Contract No. DA 36-0390SC-90742

Distribution List (Cont'd)	No. of Copies
Director U. S. Naval Research Laboratory ATTN: Code 2027 Washington 25, D. C.	1
Commanding Officer and Director U. S. Navy Electronics Laboratory SanDiego 52, California	1
Aeronautical Systems Division ATTN: Technical Library Wright-Patterson Air Force Base, Ohio	1
Hq., Electronic Systems Division ATTN: Technical Library L. G. Hanscom Field Bedford, Mass.	1
Rome Air Development Center ATTN: RAWIO Griffiss Air Force Base, New York	1
Director, Fort Monmouth Office USA Communication and Electronics Combat Development Agency Fort Monmouth, New Jersey	1
Office, Assistant Chief of Staff for Intelligence Department of the Army ATTN: Chief, Research and Development Branch Washington 25, D. C.	1
U. S. Navy Photographic Interpretation Center ATTN: Technical Library Suitland, Maryland	1
Commanding General USA Combat Developments Command ATTN: CDCMR-E Fort Belvoir, Virginia	1

Contract No. DA 36-039-SC-90742

Distribution List (Cont'd)

No.

USA Signal Missile Support Agency Office
U. S. Army Electronics
Research and Development Laboratory
Fort Monmouth, New Jersey

Marine Corps Liaison Office
U. S. Army Electronics
Research and Development Laboratory
Fort Monmouth, New Jersey

Corps of Engineers Liaison Office
U. S. Army Electronics
Research and Development Laboratory
Fort Monmouth, New Jersey

AFSC Scientific/Technical Liaison Office
U. S. Army Electronics
Research and Development Laboratory
Fort Monmouth, New Jersey

Commanding Officer
U. S. Army Liaison Group
Project Michigan
P. O. Box 618
Ann Arbor, Michigan

Commanding Officer
U. S. Army Electronics
Research and Development Laboratory
Fort Monmouth, New Jersey
ATTN: SELRA/DR

ATTN: SELRA/ADT (Tech Doc Center)

ATTN: SELRA/TNR

ATTN: SELRA/SAP

ATTN: SELRA/SA

ATTN: SELRA/SES

ATTN: Logistics Division, ATTN: SELRA/SSA, Mr. L. Dennis

ATTN: SELRA/ADO (FU #3 Evans Area)